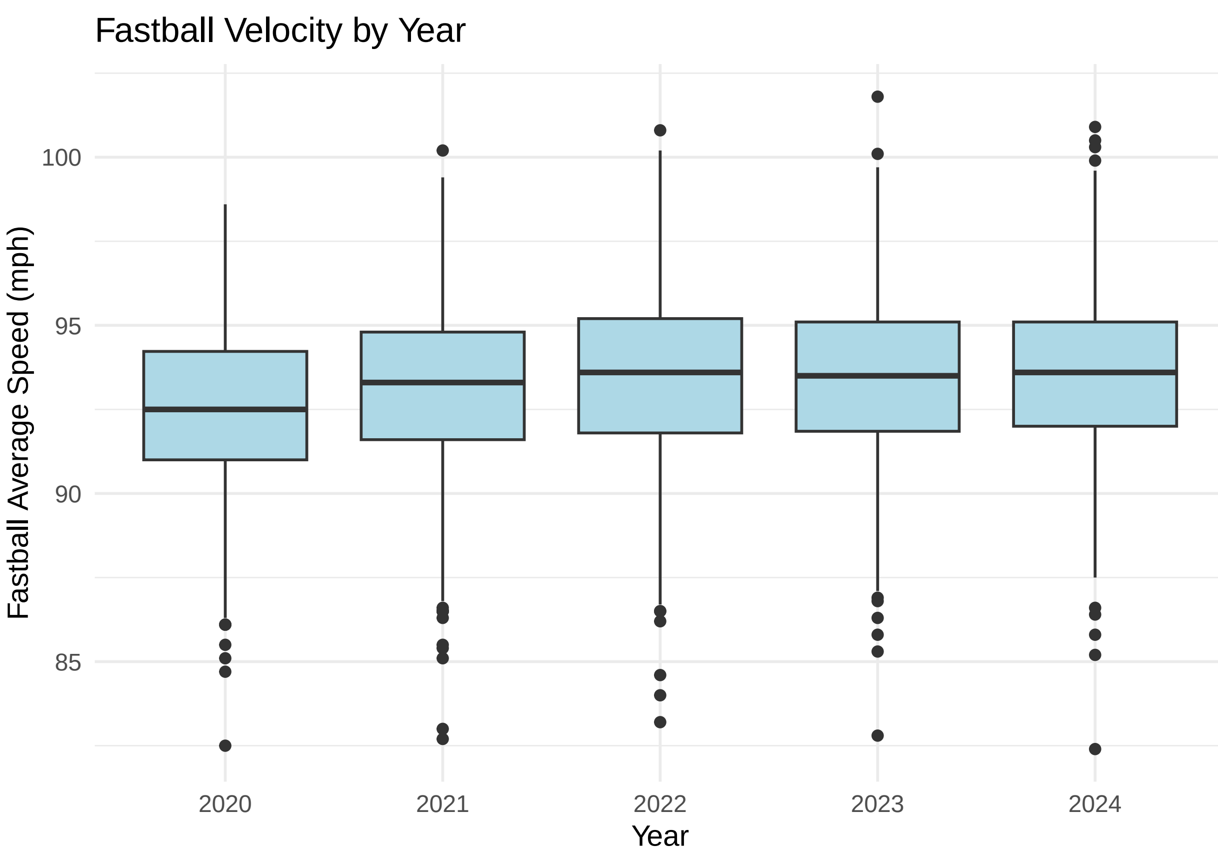
2024 Fastball Velocity Analysis

Jake Burns

2024-10-31

Fastball Velocities From the Past 5 Seasons

```
D = read.csv("Past5Pitchers.csv")
ggplot(D, aes(x = factor(year), y = fastball_avg_speed)) +
  geom_boxplot(fill = "lightblue") +
  labs(title = "Fastball Velocity by Year",
       x = "Year",
       y = "Fastball Average Speed (mph)") +
    theme_minimal()
```

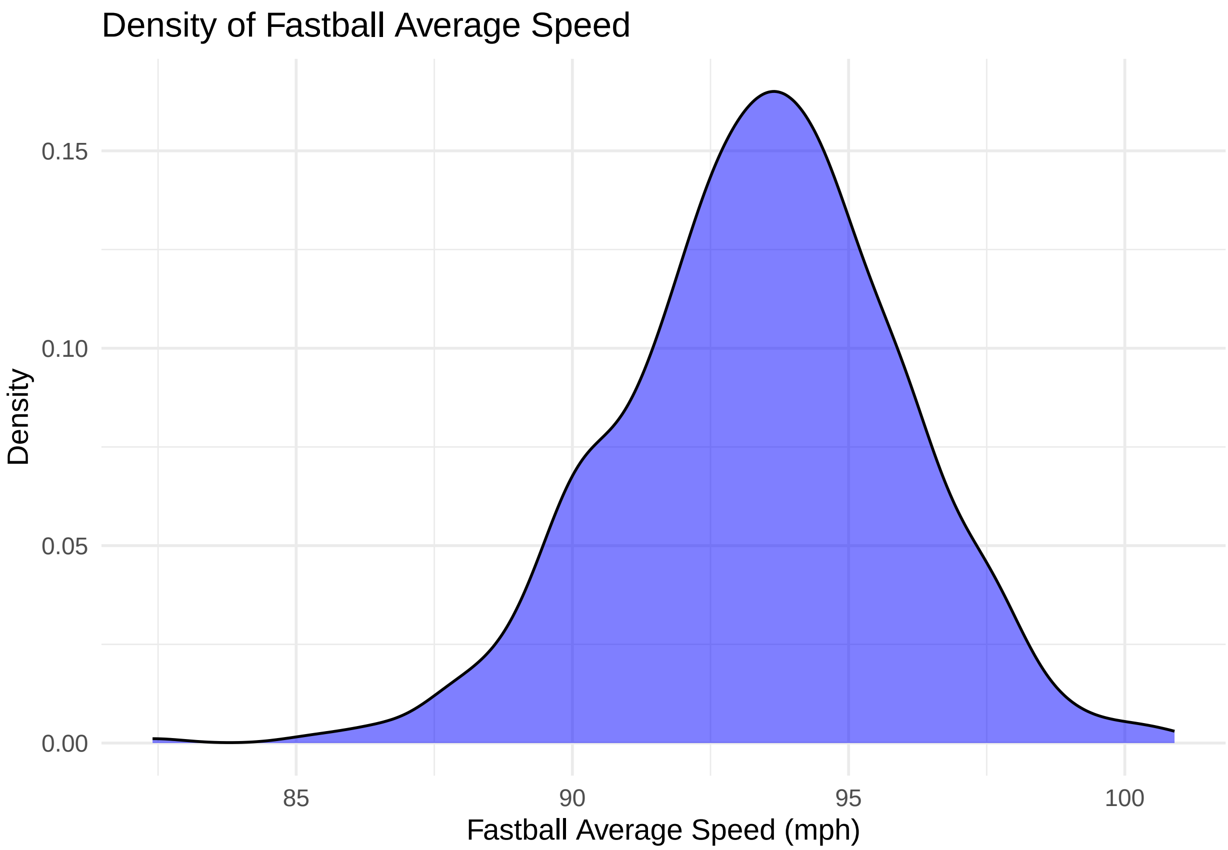


We can see from these boxplots that the average velocity of fastballs in the MLB has been slighly increasing over the years, with more outliers continuing to occur at the top. This shows that velcoity is definitely something that teams and players are valuing.

Density Plot of Fastball Velo Distribution

```
ggplot(P, aes(x = fastball_avg_speed)) +
  geom_density(fill = "blue", alpha = 0.5) +
```

```
  labs(title = "Density of Fastball Average Speed",
       x = "Fastball Average Speed (mph)",
       y = "Density") +
  theme_minimal()
```



99th Percentile Fastball Velocity

```
percentile_99_velo = quantile(P$fastball_avg_speed, 0.99)

P %>%
  select(last_name..first_name, fastball_avg_speed) %>%
  filter(fastball_avg_speed > percentile_99_velo) %>%
  arrange(desc(fastball_avg_speed))
```

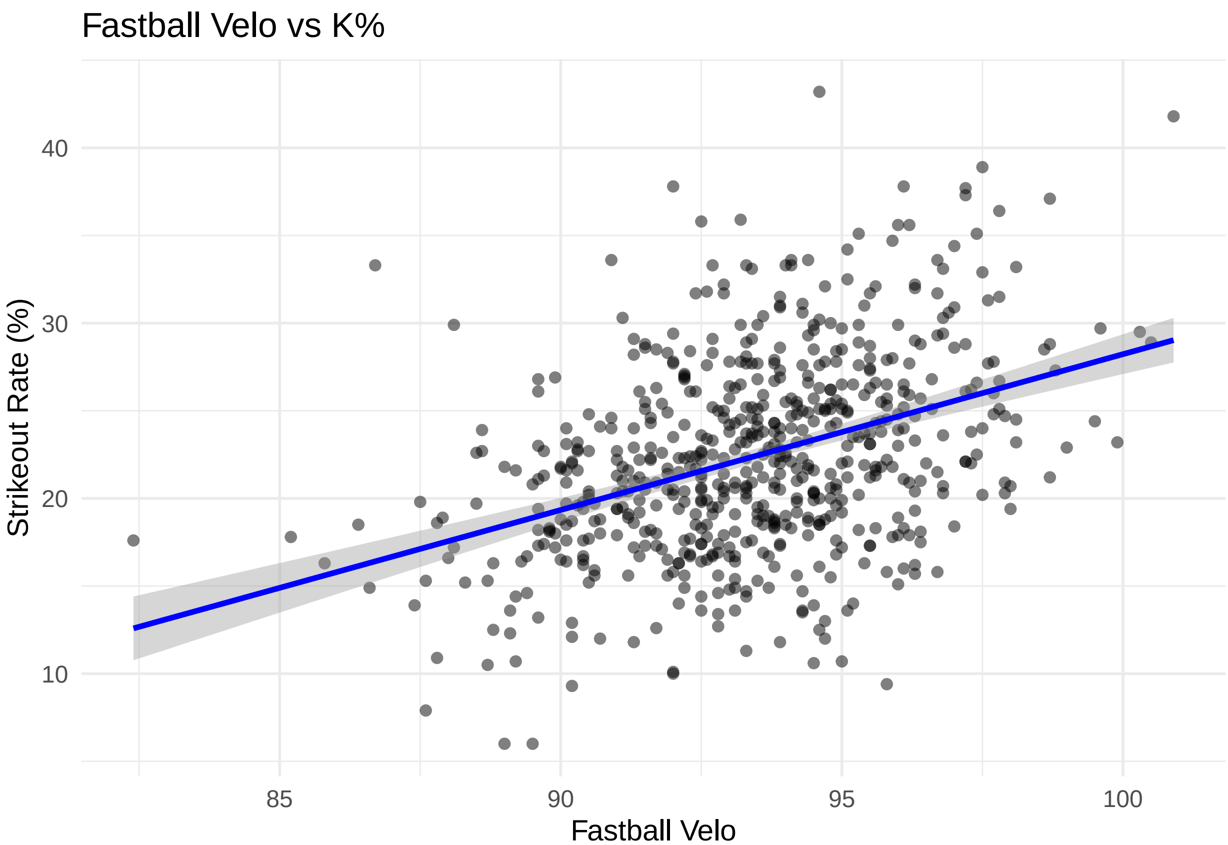
```
##   last_name..first_name fastball_avg_speed
## 1         Miller, Mason              100.9
## 2          Duran, Jhoan              100.5
## 3      Martinez, Justin              100.3
## 4            Joyce, Ben               99.9
## 5         Helsley, Ryan               99.6
## 6       Clase, Emmanuel               99.5
## 7        Suarez, Robert               99.0
```

How did velocity affect K%

```
ggplot(P, aes(x = fastball_avg_speed, y = k_percent)) +
  geom_point(alpha = 0.5) +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Fastball Velo vs K%",
```

```r
       x = "Fastball Velo",
       y = "Strikeout Rate (%)") +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

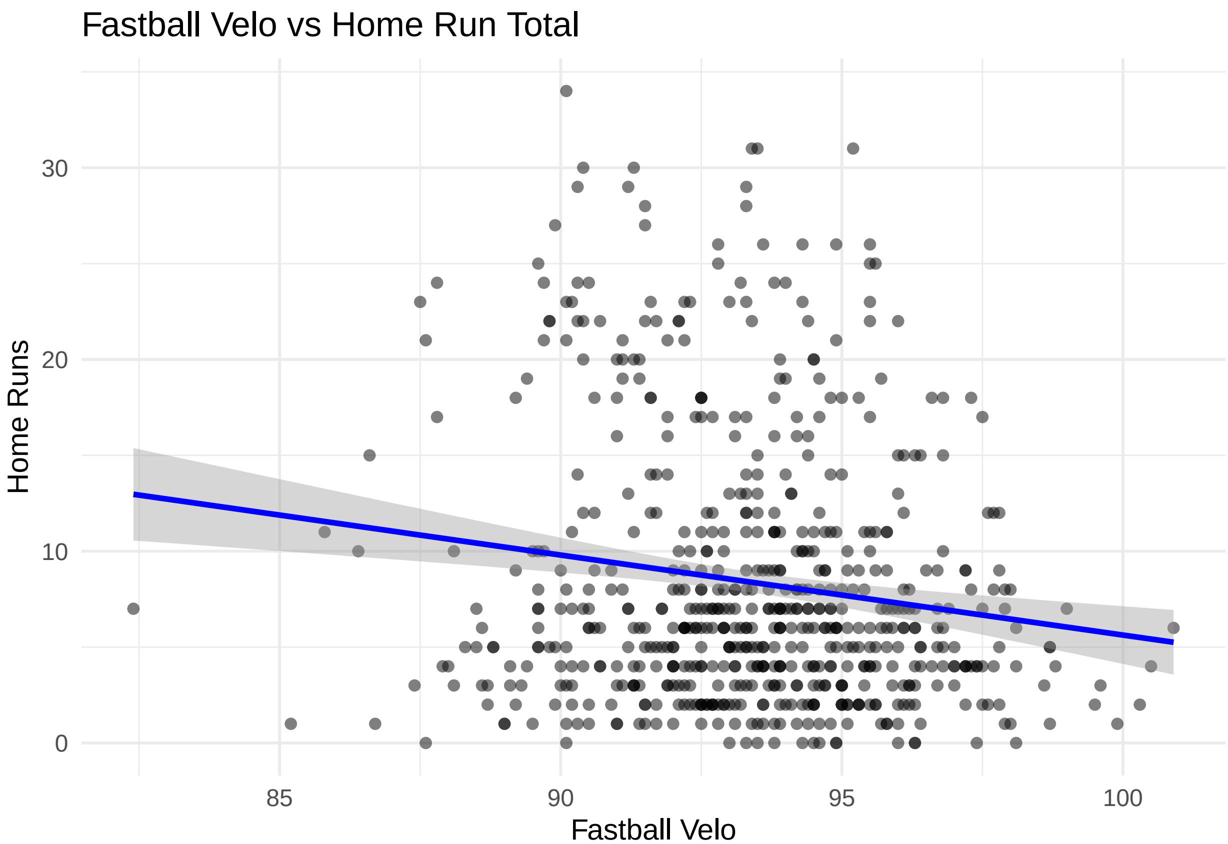


When Analyzing this graph it is clear that the increase in Fastballs Velocity is positively correlated with Strikeout Rates during the 2024 season. Many pitchers in the league are beginning to ramp up their velocity as it has to shown to be tougher for hitters to catch up to.

How did Velocity affect the total number of Home Runs given up

```r
ggplot(P, aes(x = fastball_avg_speed, y = home_run)) +
  geom_point(alpha = 0.5) +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Fastball Velo vs Home Run Total",
       x = "Fastball Velo",
       y = "Home Runs") +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

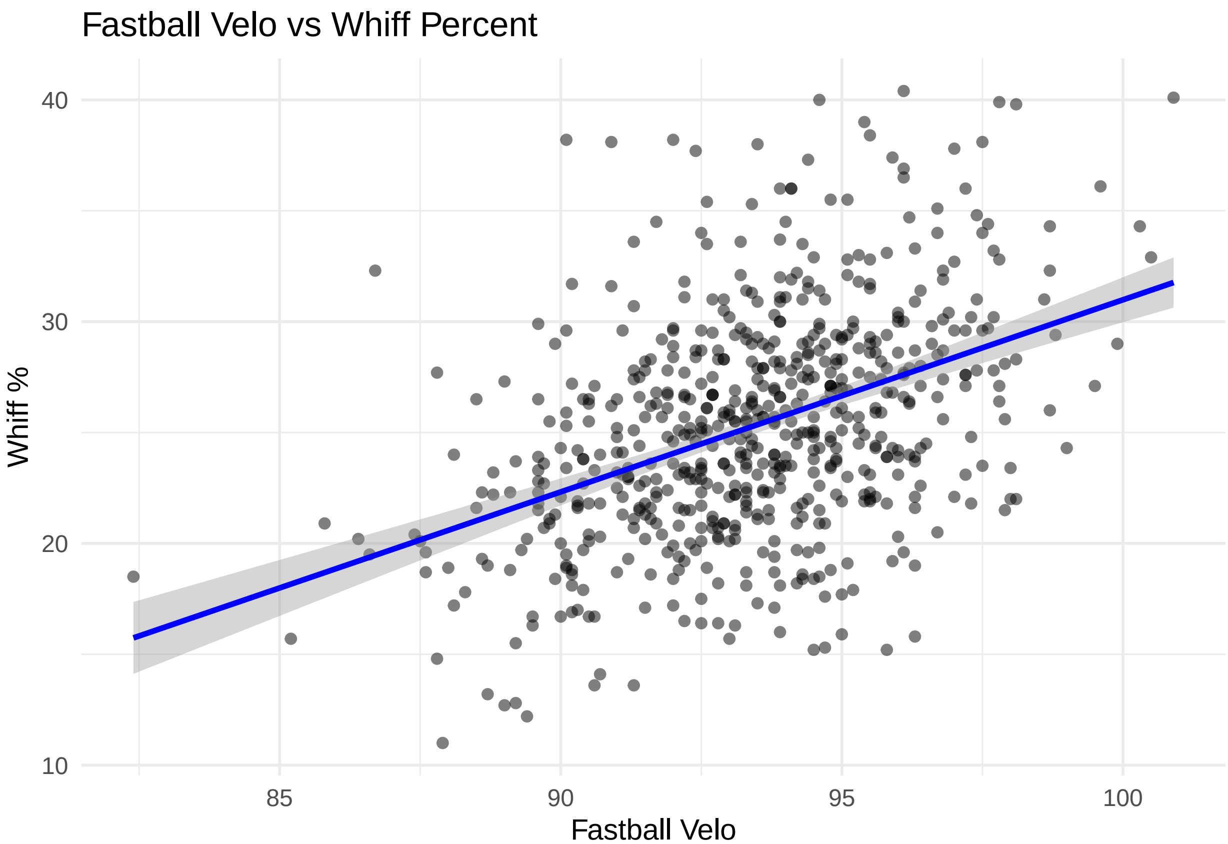


When looking at the scatterplot there is no definite correlation between the velocity and the number of home runs given up in the 2024 season. There may be a slight negative relationship between the two, showing higher velocity may have played a bit of a role in limiting home runs

How does velocity affect whiff percent

```r
ggplot(P, aes(x = fastball_avg_speed, y = whiff_percent)) +
  geom_point(alpha = 0.5) +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Fastball Velo vs Whiff Percent",
       x = "Fastball Velo",
       y = "Whiff %") +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

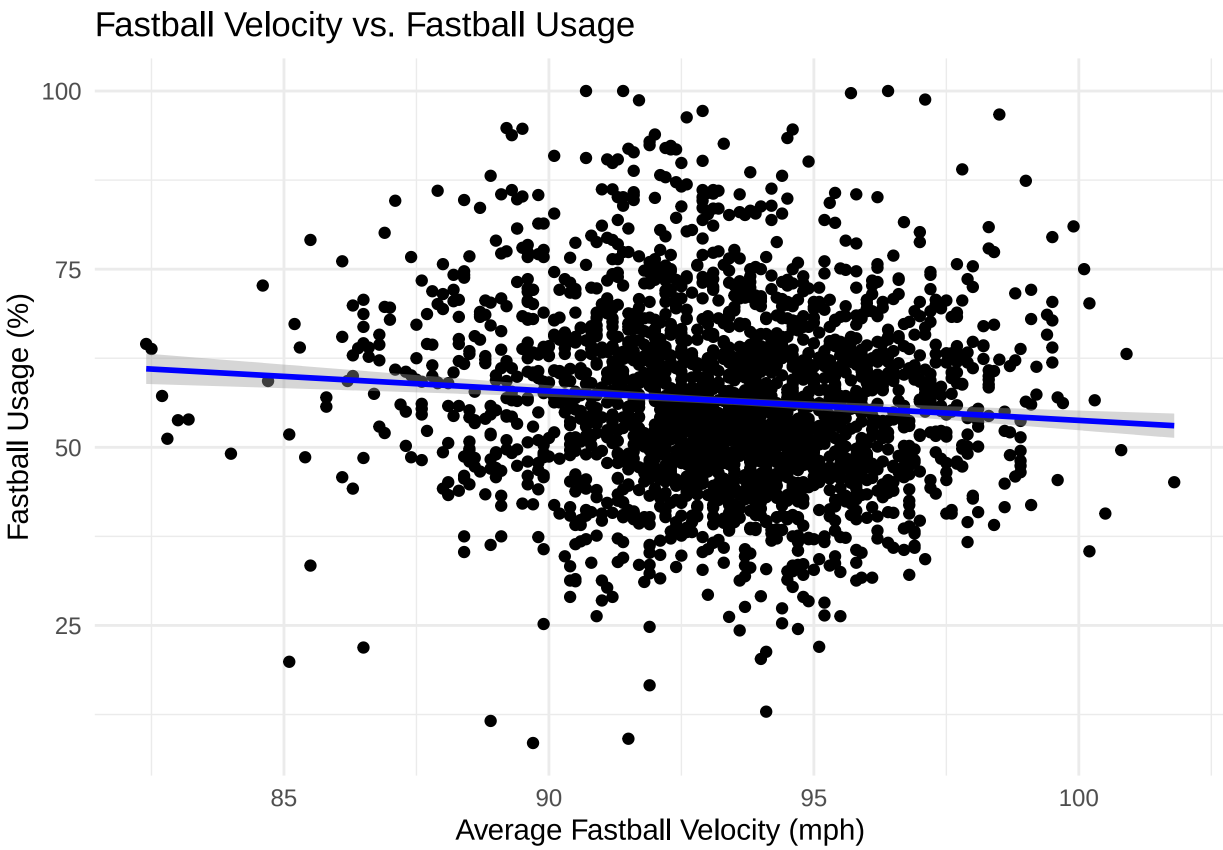


When analyzing the scatterplot here, we can see that the fastball velocity has a positive correlation with whiff rate. This makes sense as the pitches are coming in faster, making it more likley that the batters are going to swing and miss more often. While it is not the only metric that accounts for whiff%, it definitely plays a role.

Do Pitchers who throw harder use their fastball more often?

```r
ggplot(D, aes(x = fastball_avg_speed, y = n_fastball_formatted)) +
  geom_point() +
  geom_smooth(method = "lm", color = "blue") +
  labs(title = "Fastball Velocity vs. Fastball Usage",
       x = "Average Fastball Velocity (mph)",
       y = "Fastball Usage (%)") +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

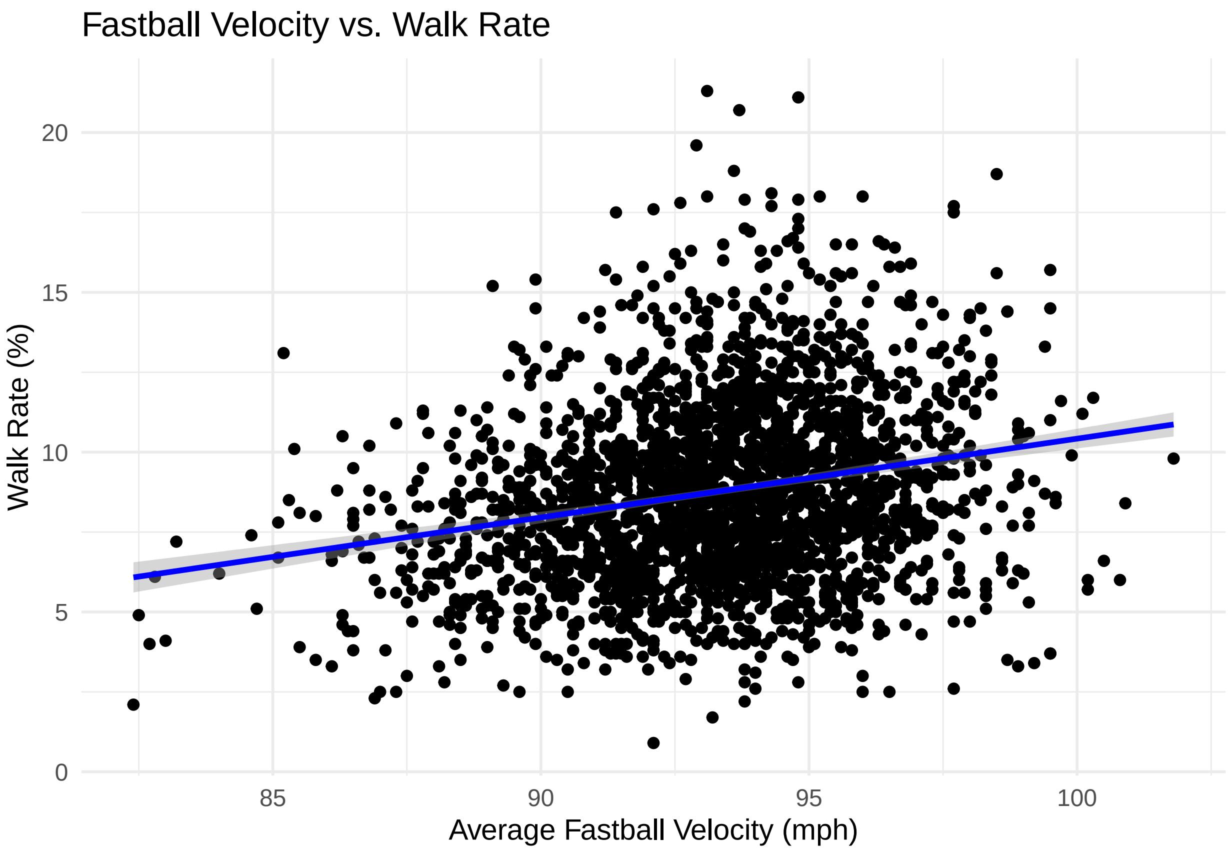


Based on the scatterplot the fastball velocity did not play a role in how often pitchers through the pitch in the 2024 season

Do pitchers walk more batters when they throw harder

```
ggplot(D, aes(x = fastball_avg_speed, y = bb_percent)) +
  geom_point() +
  geom_smooth(method = "lm", color = "blue") +
  labs(title = "Fastball Velocity vs. Walk Rate",
       x = "Average Fastball Velocity (mph)",
       y = "Walk Rate (%)") +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



Based on the scatterplot, there is a bit of a positive relationship between fastball velocity and the BB%. As the pitch velocity rises the walk rate seems to sort of rise with it as well

Does certain pitch velocity ranges lead to lower xwOBA and xSLG?

```
P = P %>%
  mutate(
    velocity_group = case_when(
      fastball_avg_speed >= 99 ~ "99+",
      fastball_avg_speed >= 95 & fastball_avg_speed < 99 ~ "95-98",
      fastball_avg_speed >= 90 & fastball_avg_speed < 95 ~ "90-94",
      fastball_avg_speed < 90 ~ "89 or Less"
    )
  )

xwoba_summary = P %>%
  group_by(velocity_group) %>%
  summarise(avg_xwoba = mean(xwoba, na.rm = TRUE)) %>%
  arrange(desc(velocity_group))

xwOBA_plot = ggplot(xwoba_summary, aes(x=velocity_group, y= avg_xwoba, fill= velocity_group)) +
  geom_bar(stat = "identity") +
  labs(title = "Average xwOBA by Fastball Velocity", x = "Fastball Velocity Group (MPH)", y= "Average xw
  theme_minimal() +
  scale_fill_brewer(palette = "Set2") +
  theme(legend.position = "none")
```

```
xslg_summary = P %>%
  group_by(velocity_group) %>%
```

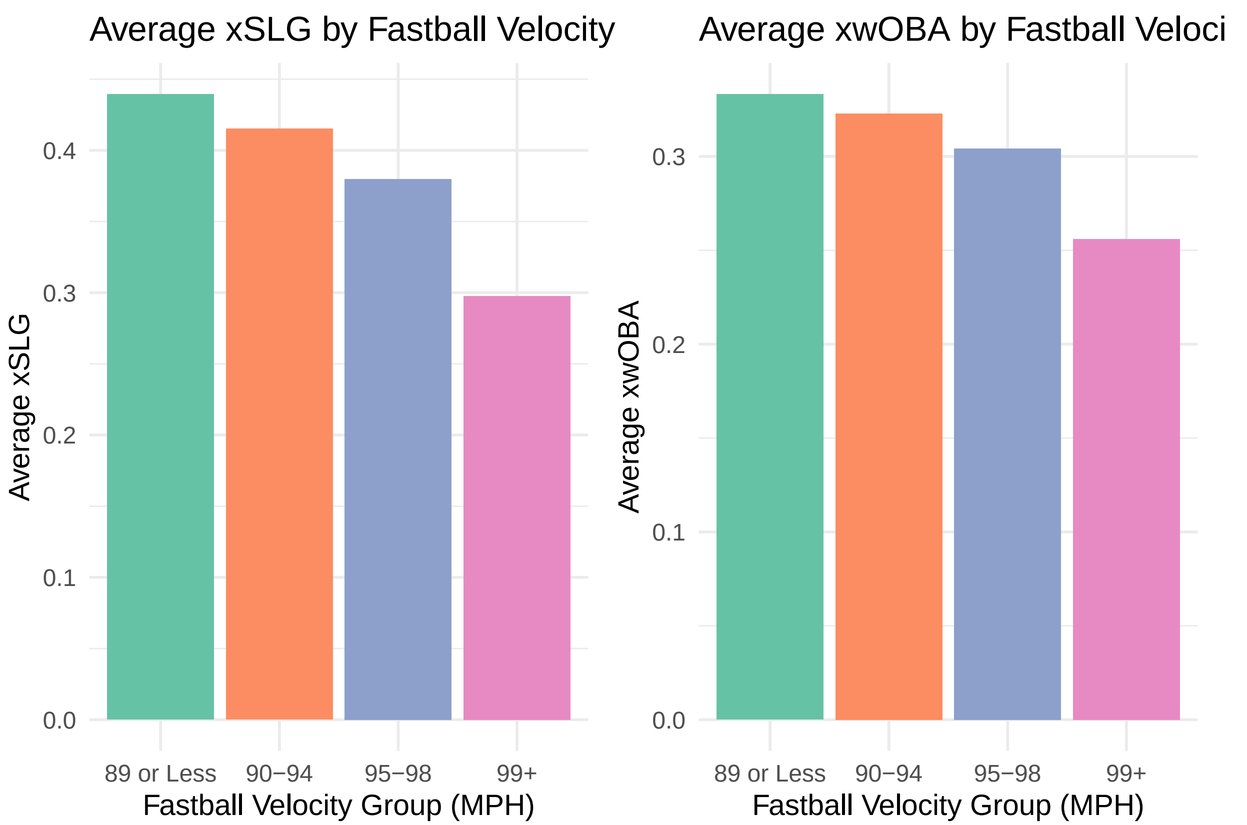
```r
  summarise(avg_xslg = mean(xslg, na.rm = TRUE)) %>%
  arrange(desc(velocity_group))

xslg_plot = ggplot(xslg_summary, aes(x=velocity_group, y= avg_xslg, fill= velocity_group)) +
  geom_bar(stat = "identity") +
  labs(title = "Average xSLG by Fastball Velocity", x = "Fastball Velocity Group (MPH)", y= "Average xSl
  theme_minimal() +
  scale_fill_brewer(palette = "Set2") +
  theme(legend.position = "none")
```

```r
library(patchwork)
```

```
## Warning: package 'patchwork' was built under R version 4.3.3
```

```r
xslg_plot + xwOBA_plot
```



As we can see with these two bars graph, as the average pitchers velocity moves up in velocity group ranges then we see a clear decline in both xSLG and xWOBA. While this may not be the only factor, how hard pitchers were throwing 2024 definitely played a role in reducing some of the most important expected statistics in hitting.

Conclusion

There are many different things that can be analyzed when it comes to pitchers in the MLB, but the velocity of the pitches is always a main thing people ask about. This project breaks down some of the statistics from the 2024 season to show how different velocities faired when coming down to pitcher performance and expected results from the batter they were facing