PitcherProject

Jake Burns

2024-10-31

```r
getwd()
```

```
## [1] "/Users/jakeburns/Desktop/PersonalProjects"
```

```r
Pitchers =read.csv("2024Pitch.csv")

library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(caret)
```

```
## Loading required package: lattice
```

#Analyzing Horinzontal Movement From Different Arm Angles

Model of Horizontal Break Comapred to Arm Angle

```r
model_FHB <- lm(fast_avg_horizontal_break ~ arm_angle, data = Pitchers)
summary(model_FHB)
```

```
##
## Call:
## lm(formula = fast_avg_horizontal_break ~ arm_angle, data = Pitchers)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -24.1432  -2.4904   0.2678   2.4707   8.5478
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 14.76537    0.50829   29.05   <2e-16 ***
## arm_angle   -0.16190    0.01274  -12.71   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

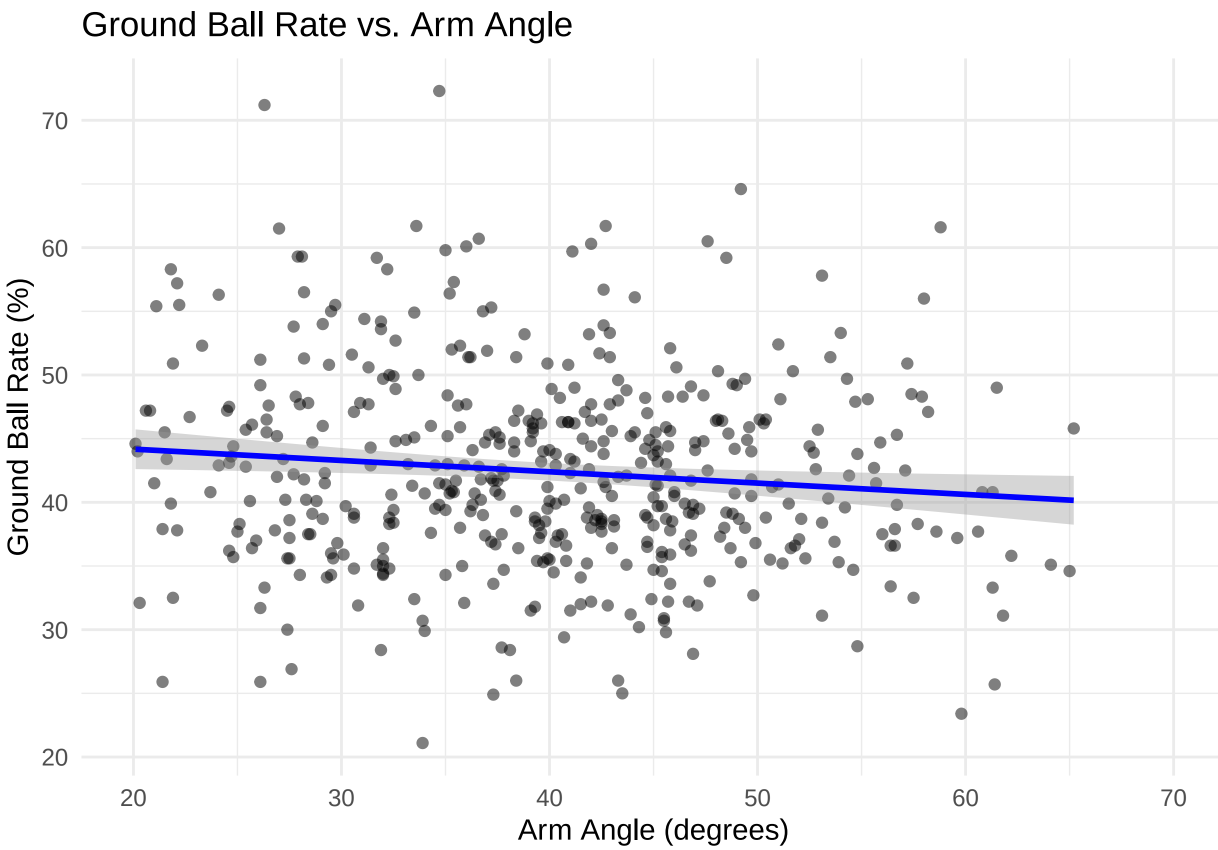
```
##
## Residual standard error: 3.835 on 515 degrees of freedom
## Multiple R-squared:  0.2387, Adjusted R-squared:  0.2372
## F-statistic: 161.4 on 1 and 515 DF,  p-value: < 2.2e-16
```

The analysis shows a significant negative relationship between arm_angle and fast_avg_horizontal_break, suggesting that as arm angle increases, the horizontal break on fastballs tends to decrease. The relatively high R-squared value indicates that arm angle explains a notable portion of the variability in horizontal break. This analysis suggests that pitchers may need to consider their arm angle when aiming to optimize the horizontal movement on their fastballs. Further investigation could examine the interactions of arm angle with other mechanical factors and their collective impact on pitch movement.

Do Ground Ball Rates Increase as a Result of Dropping Your Arm Angle?

```r
ggplot(Pitchers, aes(x = arm_angle, y = groundballs_percent)) +
  geom_point(alpha = 0.5) +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Ground Ball Rate vs. Arm Angle",
       x = "Arm Angle (degrees)",
       y = "Ground Ball Rate (%)") +
  xlim(20,70) +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
## Warning: Removed 40 rows containing non-finite values (`stat_smooth()`).
## Warning: Removed 40 rows containing missing values (`geom_point()`).
```




```r
model_gb <- lm(groundballs_percent ~ arm_angle, data = Pitchers)
summary(model_gb)
```

```
##
## Call:
## lm(formula = groundballs_percent ~ arm_angle, data = Pitchers)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -22.1313  -5.0083  -0.6784   4.5503  29.1698
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 47.51745    1.03154  46.065  < 2e-16 ***
## arm_angle   -0.12643    0.02586  -4.889 1.35e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.783 on 515 degrees of freedom
## Multiple R-squared:  0.04436,    Adjusted R-squared:  0.0425
## F-statistic:  23.9 on 1 and 515 DF,  p-value: 1.355e-06
```

The analysis indicates a significant negative relationship between arm_angle and groundballs_percent, suggesting that as the arm angle increases, the percentage of ground balls decreases. Although the R-squared value indicates that a relatively small proportion of the variance in ground ball percentage is explained by

#arm angle, the statistical significance of the coefficient indicates that arm angle is a relevant predictor. Further research could explore additional variables that might influence ground ball rates or the interactions between arm angle and other pitching mechanics.

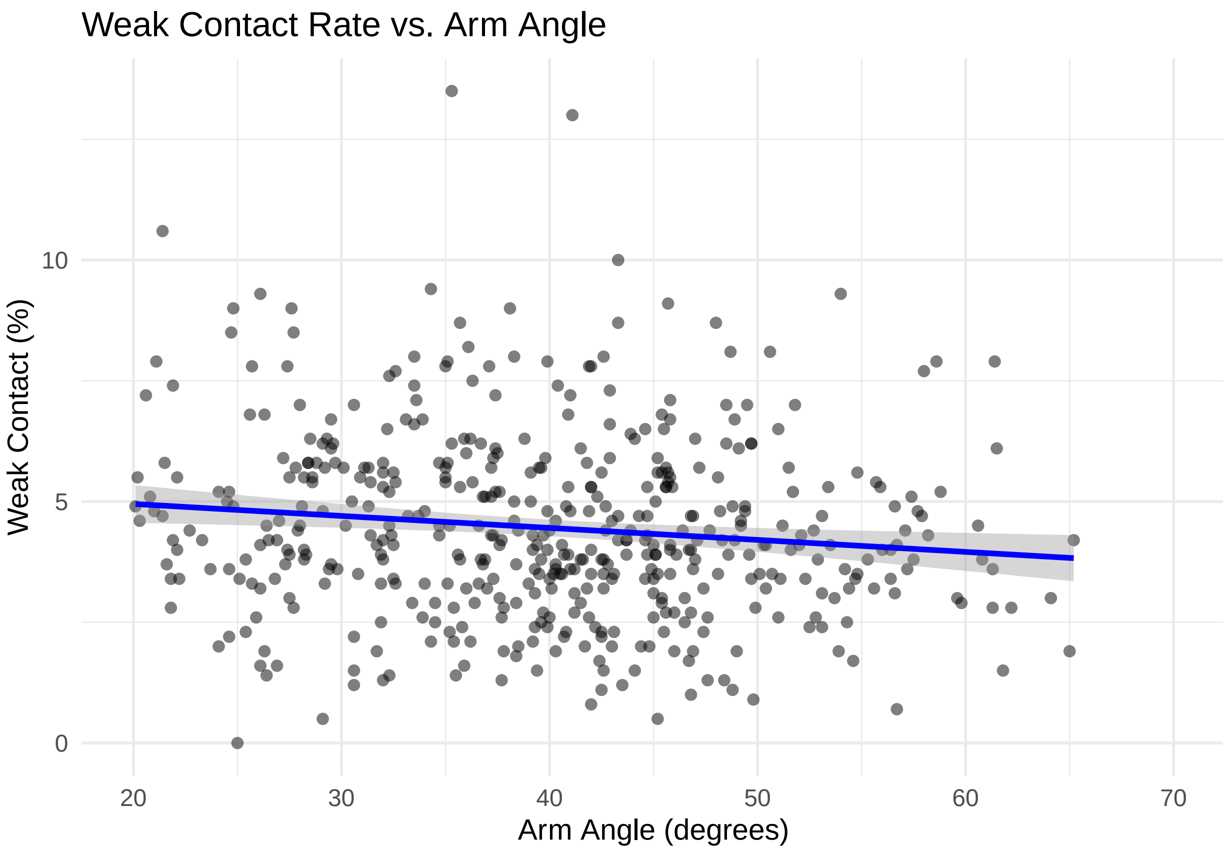
Does Dropping Arm Angle Create More Weak Contact?

```r
ggplot(Pitchers, aes(x = arm_angle, y = poorlyweak_percent)) +
  geom_point(alpha = 0.5) +
  geom_smooth(method = "lm", col = "blue") +
  labs(title = "Weak Contact Rate vs. Arm Angle",
       x = "Arm Angle (degrees)",
       y = "Weak Contact (%)") +
  xlim(20,70) +
  theme_minimal()
```

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

```
## Warning: Removed 40 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 40 rows containing missing values (`geom_point()`).
```




```r
model_weakcontact <- lm(poorlyweak_percent ~ arm_angle, data = Pitchers)
summary(model_weakcontact)
```

```
##
## Call:
## lm(formula = poorlyweak_percent ~ arm_angle, data = Pitchers)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.857 -1.245 -0.222  1.080  8.910
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.505921   0.255486  21.551  < 2e-16 ***
## arm_angle   -0.025956   0.006405  -4.053 5.85e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.928 on 515 degrees of freedom
## Multiple R-squared:  0.0309, Adjusted R-squared:  0.02902
## F-statistic: 16.42 on 1 and 515 DF,  p-value: 5.848e-05
```

In summary, while arm_angle is statistically significant in predicting poorlyweak_percent, the low R^2 value indicates that the model explains only a small fraction of the variability in poorly hit balls. This suggests that there are likely other influential factors or predictors that could be included to improve the model's explanatory power. The significant negative relationship implies that adjustments in arm_angle may help reduce poorly hit balls, but further exploration of additional variables may provide a more comprehensive understanding of the factors affecting poorlyweak_percent.