Duke Defensive Stats 2023

Packages

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.3     v readr     2.1.4
v forcats   1.0.0     v stringr   1.5.0
v ggplot2   3.4.3     v tibble    3.2.1
v lubridate 1.9.2     v tidyr     1.3.0
v purrr     1.0.2
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becor
```

```
library(tidymodels)
```

```
-- Attaching packages --------------------------------------- tidymodels 1.1.1 --
v broom        1.0.5     v rsample      1.2.0
v dials        1.2.0     v tune         1.1.2
v infer        1.0.4     v workflows    1.1.3
v modeldata    1.2.0     v workflowsets 1.0.1
v parsnip      1.1.1     v yardstick    1.2.0
v recipes      1.0.8
-- Conflicts ------------------------------------------ tidymodels_conflicts() --
x scales::discard() masks purrr::discard()
x dplyr::filter()   masks stats::filter()
x recipes::fixed()  masks stringr::fixed()
x dplyr::lag()      masks stats::lag()
```

```
x yardstick::spec() masks readr::spec()
x recipes::step()   masks stats::step()
* Dig deeper into tidy modeling with R at https://www.tmwr.org
```

```
duke_stats <- read_csv("data/Duke Defense Stats - DukeData.csv")
```

```
Rows: 156 Columns: 11
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (4): OppName, Surface, Site, Type
dbl (7): FPI, FPI_diff, Month, Year, Count, Yards, TD Gained

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
duke_stats
```

```
# A tibble: 156 x 11
   OppName              FPI FPI_diff Surface Month  Year Site  Type  Count Yards
   <chr>              <dbl>    <dbl> <chr>   <dbl> <dbl> <chr> <chr> <dbl> <dbl>
 1 Clemson             13.8      4.8 Grass       9  2023 Home  Inte~     1     0
 2 Lafayette           NA       NA   Grass       9  2023 Home  Inte~     2     0
 3 Northwestern         0.8     -8.2 Grass       9  2023 Home  Inte~     1     5
 4 Connecticut        -15.9    -24.9 Grass       9  2023 Away  Inte~     0     0
 5 Notre Dame          20.7     11.7 Grass       9  2023 Home  Inte~     0     0
 6 North Carolina St.   6.9     -2.1 Grass      10  2023 Home  Inte~     1    24
 7 Florida St.         19.8     10.8 Grass      10  2023 Away  Inte~     1    13
 8 Louisville          11.4      2.4 Turf       10  2023 Away  Inte~     0     0
 9 Wake Forest         -1.7    -10.7 Grass      11  2023 Home  Inte~     1    -1
10 North Carolina      10.2      1.2 Turf       11  2023 Away  Inte~     1    12
# i 146 more rows
# i 1 more variable: `TD Gained` <dbl>
```

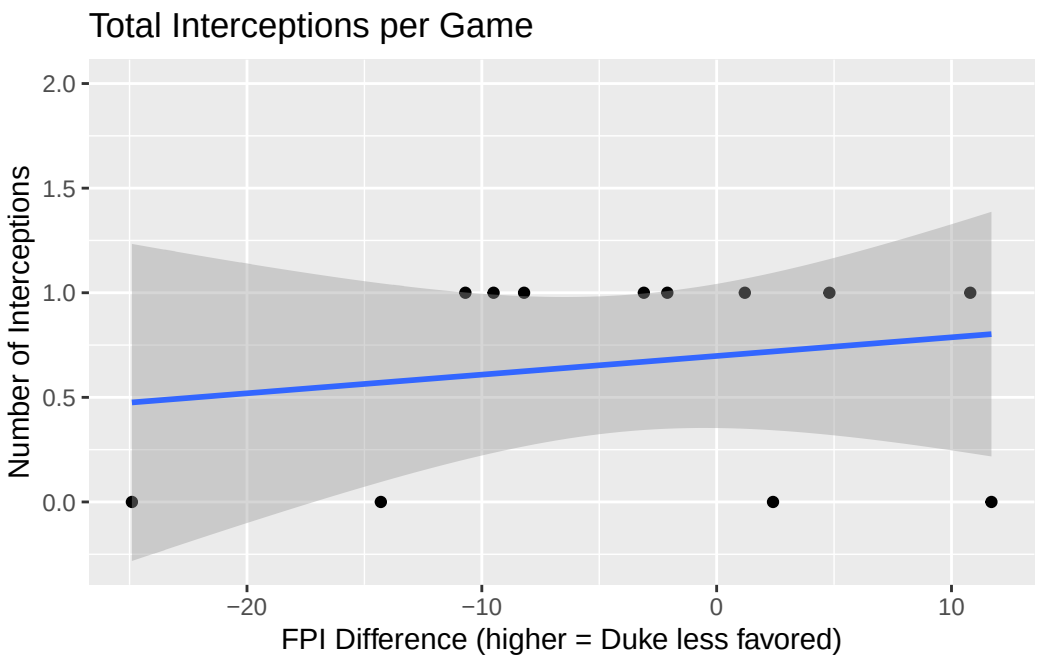
```
duke_stats |>
  filter(Type == "Interceptions") |>
  ggplot(
    aes(x = FPI_diff, y = Count)
  ) +
  geom_point() +
```

```
    geom_smooth(method = "lm", se = TRUE) +
    labs(title = "Total Interceptions per Game",
         x = "FPI Difference (higher = Duke less favored)",
         y = "Number of Interceptions")
```

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

```
Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
```

```
Warning: Removed 1 rows containing missing values (`geom_point()`).
```

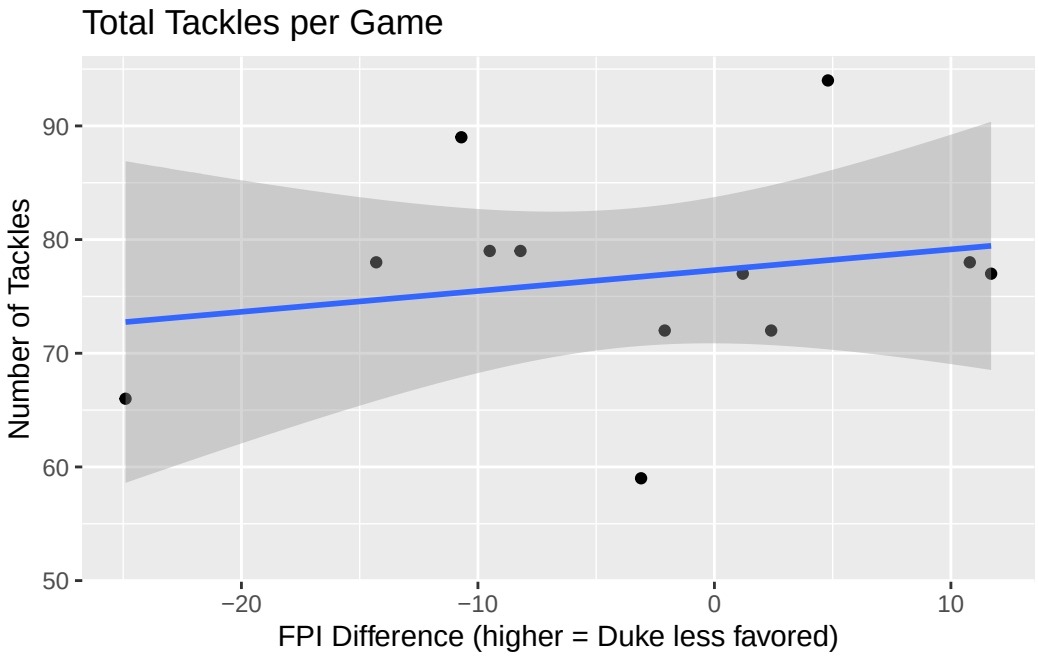



```
duke_stats |>
  filter(Type == "Tackles_Total") |>
  ggplot(
    aes(x = FPI_diff, y = Count)
  ) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  labs(title = "Total Tackles per Game",
```

```
        x = "FPI Difference (higher = Duke less favored)",
        y = "Number of Tackles")
```

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

```
Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
Removed 1 rows containing missing values (`geom_point()`).
```



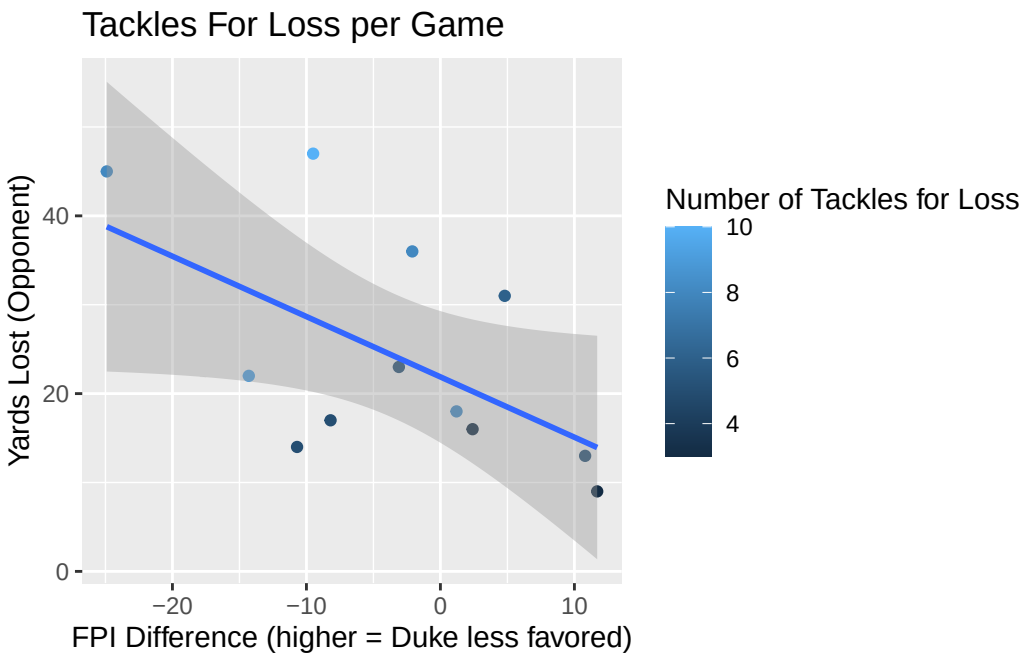

```
duke_stats |>
  filter(Type == "Tackles_for_Loss") |>
  ggplot(
    aes(x = FPI_diff, y = Yards, color = Count)
  ) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  labs(title = "Tackles For Loss per Game",
       x = "FPI Difference (higher = Duke less favored)",
       y = "Yards Lost (Opponent)",
       color = "Number of Tackles for Loss")
```

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

Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).

Warning: The following aesthetics were dropped during statistical transformation: colour
i This can happen when ggplot fails to infer the correct grouping structure in
  the data.
i Did you forget to specify a `group` aesthetic or to convert a numerical
  variable into a factor?

Warning: Removed 1 rows containing missing values (`geom_point()`).
```

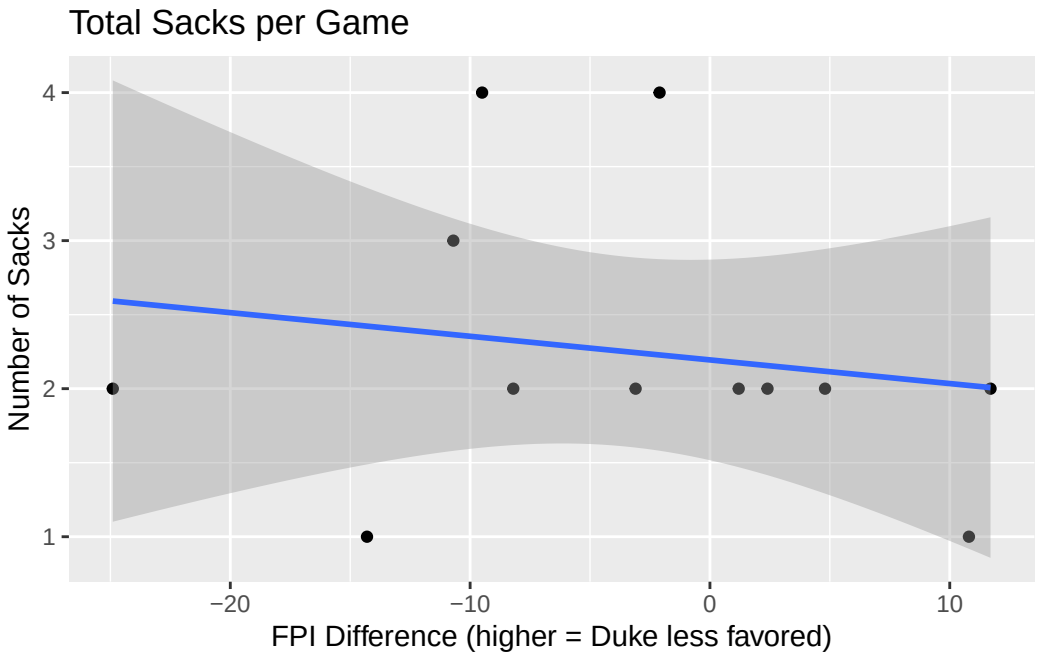



```
duke_stats |>
  filter(Type == "Sacks") |>
  ggplot(
    aes(x = FPI_diff, y = Count)
  ) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  labs(title = "Total Sacks per Game",
```

```
        x = "FPI Difference (higher = Duke less favored)",
        y = "Number of Sacks")
```

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

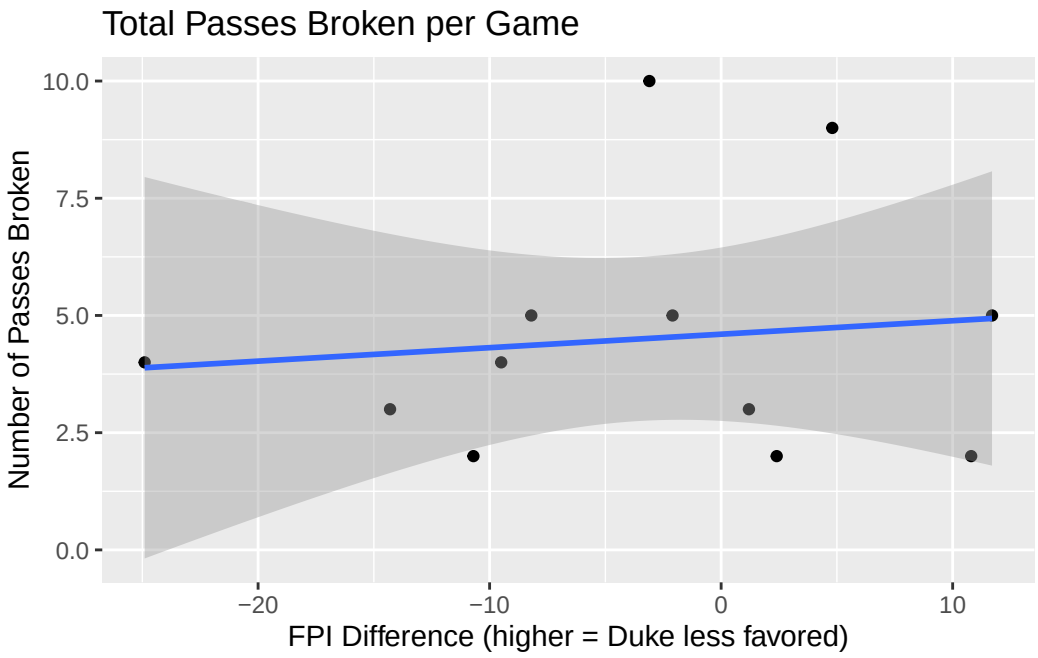
```
Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
Removed 1 rows containing missing values (`geom_point()`).
```




```
duke_stats |>
  filter(Type == "Passes_Broken") |>
  ggplot(
    aes(x = FPI_diff, y = Count)
  ) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  labs(title = "Total Passes Broken per Game",
       x = "FPI Difference (higher = Duke less favored)",
       y = "Number of Passes Broken")
```

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

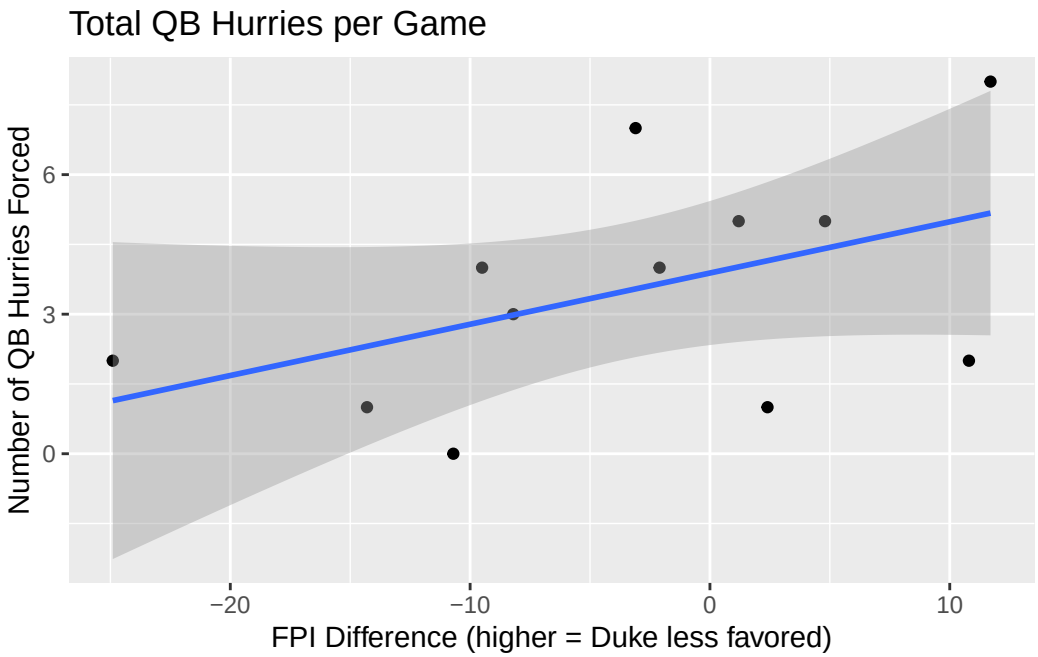
```
Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
Removed 1 rows containing missing values (`geom_point()`).
```




```
duke_stats |>
  filter(Type == "QB_Hurries") |>
  ggplot(
    aes(x = FPI_diff, y = Count)
  ) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  labs(title = "Total QB Hurries per Game",
       x = "FPI Difference (higher = Duke less favored)",
       y = "Number of QB Hurries Forced")
```

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

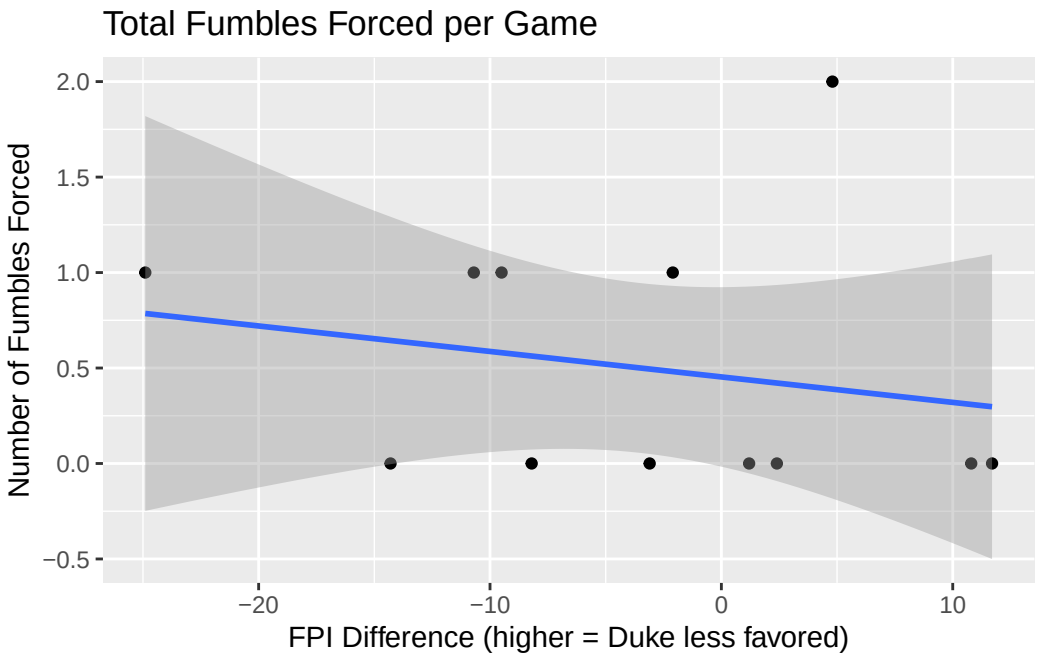
```
Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
Removed 1 rows containing missing values (`geom_point()`).
```




```
duke_stats |>
  filter(Type == "Fumbles_Forced") |>
  ggplot(
    aes(x = FPI_diff, y = Count)
  ) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  labs(title = "Total Fumbles Forced per Game",
       x = "FPI Difference (higher = Duke less favored)",
       y = "Number of Fumbles Forced")
```

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

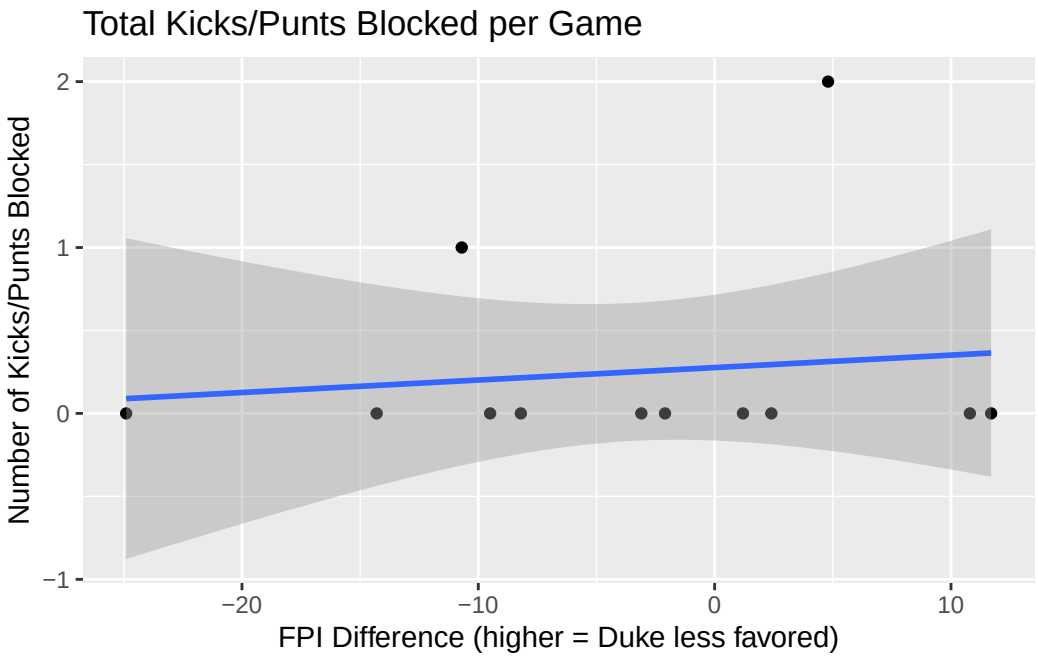
```
Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
Removed 1 rows containing missing values (`geom_point()`).
```




```
duke_stats |>
  filter(Type == "Kicks_Punts_Blocked") |>
  ggplot(
    aes(x = FPI_diff, y = Count)
  ) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  labs(title = "Total Kicks/Punts Blocked per Game",
       x = "FPI Difference (higher = Duke less favored)",
       y = "Number of Kicks/Punts Blocked")
```

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

```
Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
Removed 1 rows containing missing values (`geom_point()`).
```




```
duke_stats |>
  filter(Type == "Opp_Penalties") |>
  ggplot(
    aes(x = FPI_diff, y = Yards, color = Count)
  ) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  labs(title = "Opponent Penalty Yards per Game",
       x = "FPI Difference (higher = Duke less favored)",
       y = "Number of Opponent Penalty Yards",
       color = "Number of Opponent Penalties")
```

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

```
Warning: Removed 1 rows containing non-finite values (`stat_smooth()`).
```

```
Warning: The following aesthetics were dropped during statistical transformation: colour
i This can happen when ggplot fails to infer the correct grouping structure in
  the data.
i Did you forget to specify a `group` aesthetic or to convert a numerical
  variable into a factor?
```

```
Warning: Removed 1 rows containing missing values (`geom_point()`).
```