EPL Payroll vs Performance Analysis

Thinh Nguyen

December 1st, 2024

1 Data Summary

We analyzed payroll and performance data for English Premier League teams from the 2016–17 to 2023–24 seasons.

The dataset was imported and summarized using the following code:

```
setwd("~/University/Georgia␣Tech/Courses/Fall␣2024/ECON␣2250/R␣
   Project")
ptsData <- read.csv("Raw␣data.csv")

summary(ptsData[, c("payroll", "pts")])
```

Table 1 shows the descriptive statistics. Two histograms of Points and Payrolls are also provided.

	Payroll (€)	Points
Min	17,336,600	16
1st Qu.	52,706,000	40
Median	75,168,900	49
Mean	89,700,763	53.15
3rd Qu.	129,454,568	66
Max	260,630,000	100

Table 1: Summary statistics for payroll and points

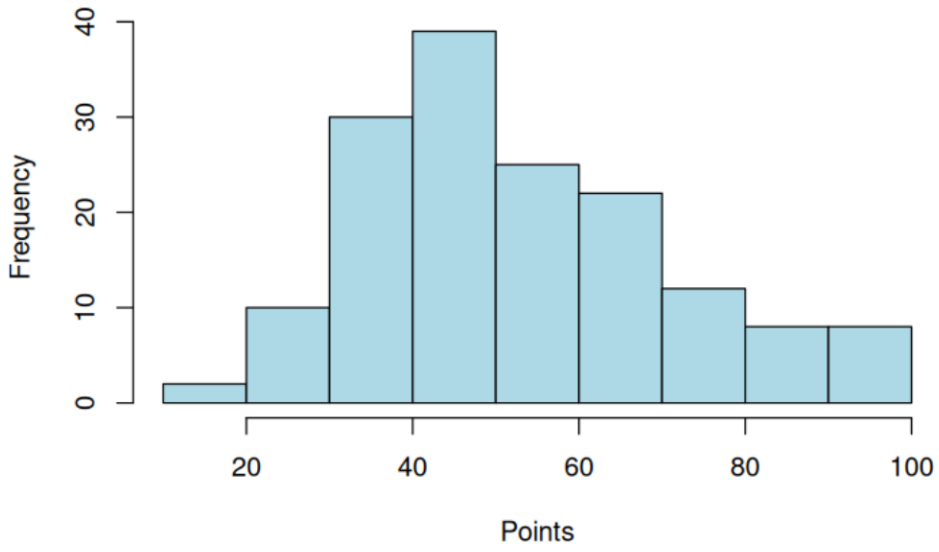

Figure 1: Histogram of Points

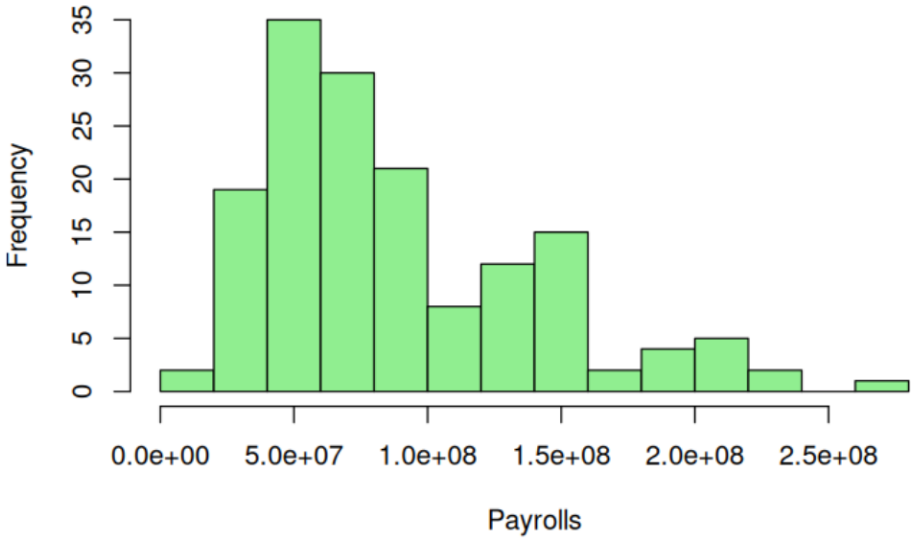

Figure 2: Histogram of Payrolls

2 Hypothesis Testing

Teams were categorized into two groups:

- High-payroll teams: payroll $>$ €100M
- Low-payroll teams: payroll $\leq$ €100M

2.1 One-Sided Test

We tested whether high-payroll teams earned more points on average:

$$H_0 : \mu_{\text{high}} \leq \mu_{\text{low}} \qquad H_a : \mu_{\text{high}} > \mu_{\text{low}}$$

The test was performed using the following code:

```
t_test_result <- t.test(high_payroll$pts, low_payroll$pts,
                        alternative = "greater", var.equal = TRUE)
t_test_result
```

Results:

- $t = 13.382$, $df = 154$, p-value $< 2.2e^{-16}$
- Mean (high-payroll) = 73.29
- Mean (low-payroll) = 43.93

Since $p < 0.05$, we reject H_0 and conclude that high-payroll teams perform better.

2.2 Two-Sided Test

We further tested whether the two means differ:

$$H_0 : \mu_{\text{high}} = \mu_{\text{low}} \qquad H_a : \mu_{\text{high}} \neq \mu_{\text{low}}$$

The test was performed using the following code:

```
t_test_result2 <- t.test(high_payroll$pts, low_payroll$pts,
                         alternative = "two.sided", var.equal = TRUE
                            )
t_test_result2
```

Results:

- $t = 13.382$, $df = 154$, p-value $< 2.2e^{-16}$
- 95% CI (Confidence Interval) for difference in means: [25.03, 33.69]
- Estimated difference in means: 29.36

We conclude that high-payroll teams average between 25–34 more points per season than low-payroll teams.

3 Regression Analysis

A simple linear regression was performed with payroll as the independent variable and points as the dependent variable:

$$\hat{Y} = 30.44 + (2.53 \times 10^{-7}) \cdot X$$

where X is team payroll. The positive coefficient indicates that higher payroll is associated with higher points.

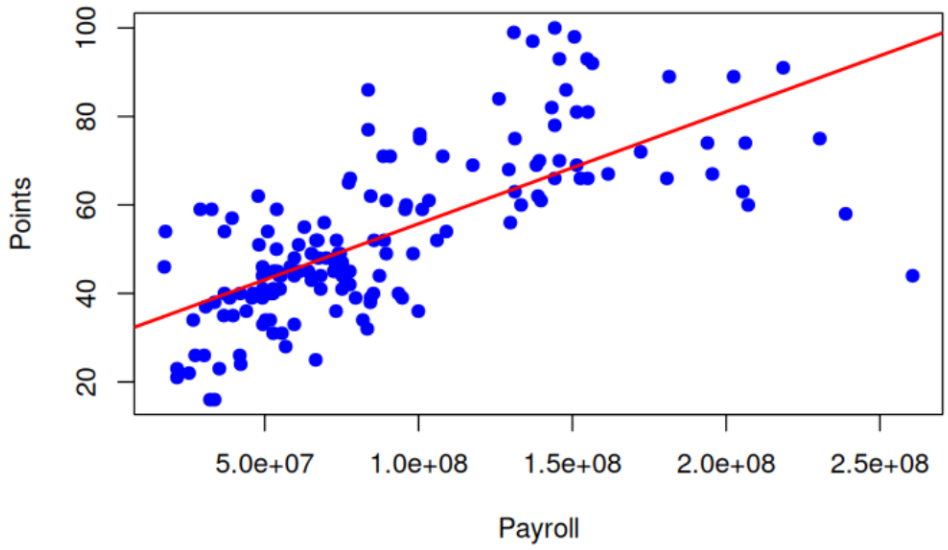

Figure 3: Scatter Plot: Payroll vs. Points with Regression Line

4 Conclusion

The results show that payroll is a strong predictor of performance. High-payroll teams earn significantly more points, averaging nearly 30 points more than low-payroll teams per season. Financial investment is therefore a key driver of success in the English Premier League.