

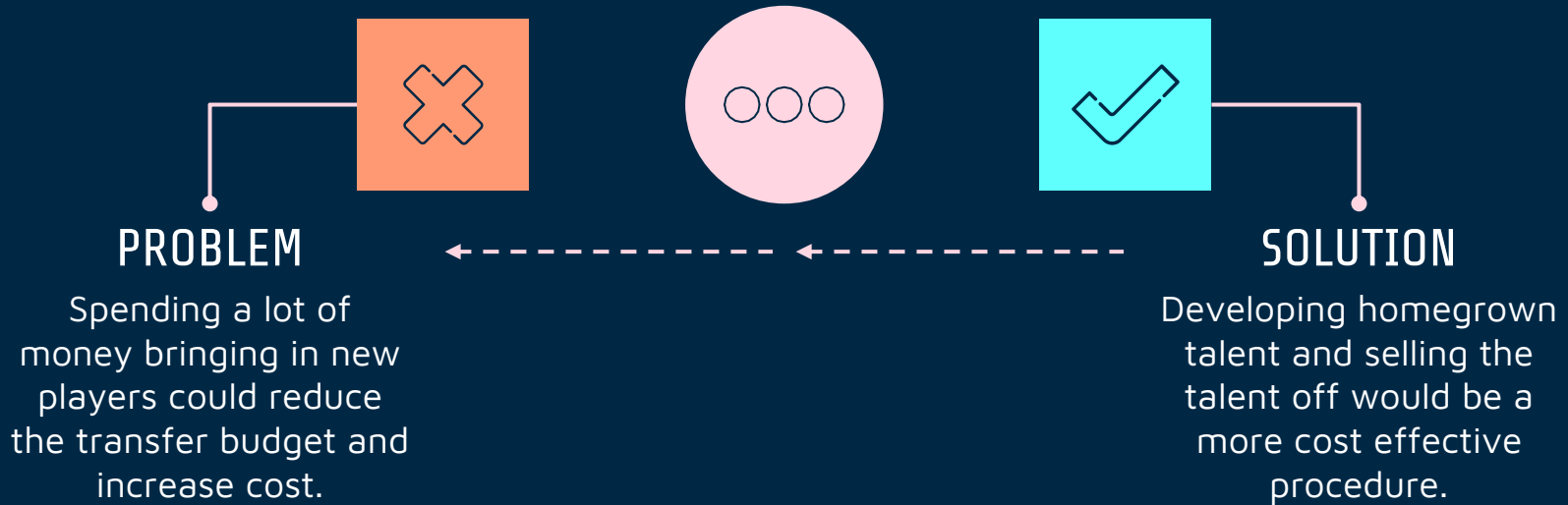
The background is a dark blue gradient. It is decorated with various geometric elements: small squares in teal, orange, and pink, and thin white vertical lines of varying lengths. Some squares are solid, while others are outlines. The lines are positioned at different heights and widths, creating a dynamic, abstract pattern.

# THE MYSTERY OF PLAYER VALUE

By, Ikenna Atum

# INTRODUCTION

## THE PROCESS



# INTRODUCTION

- The response variable for the analysis:
  - Value\_eur
- The predictors for the analysis include:
  - Wage\_eur
  - Release\_Clause\_Eur
  - Pace
  - Shooting
  - Passing
  - Defending



# Data Sources

- Fifa 21 Data from Kaggle
  - Number of columns : 134
  - Number of rows: 189444



# Hypothesis

- Null Hypothesis: There is no relationship between the dependent variable and the predictors.
  - $H_0: \beta_0 + \beta_1 + \beta_2 + \dots + \beta_k = 0$
- Alternative Hypothesis: At least one of the predictors is  $\beta_1$  is correlated with the response.
  - $H_0: \beta_0 + \beta_1 + \beta_2 + \dots + \beta_k \neq 0$

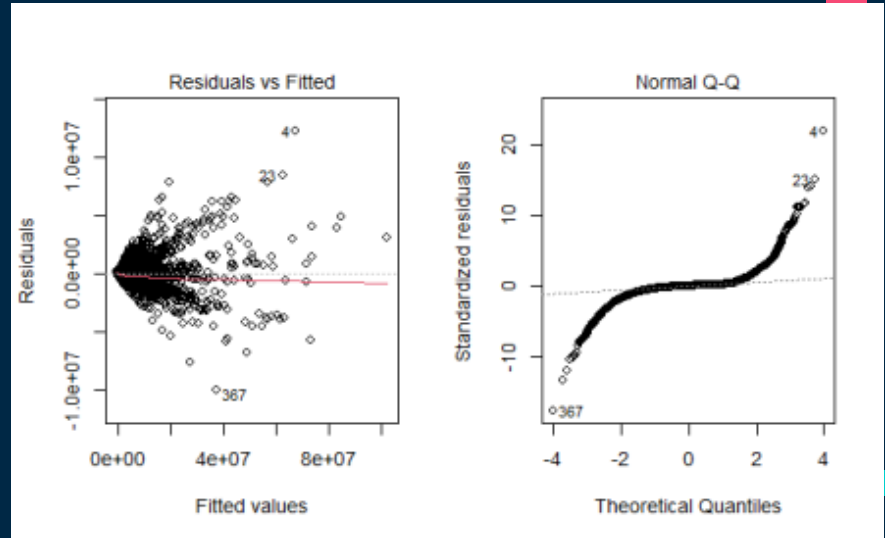


# Model

```
lm(value_eur ~ overall + wage_eur + release_clause_eur + pace + shooting +  
passing + dribbling + defending, data = Playas21)
```

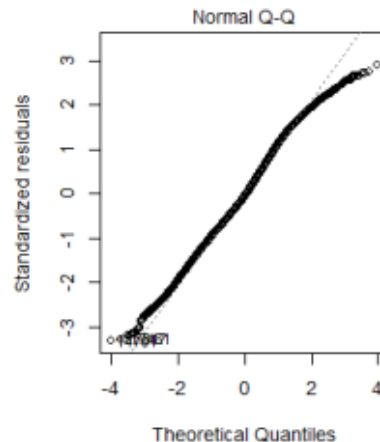
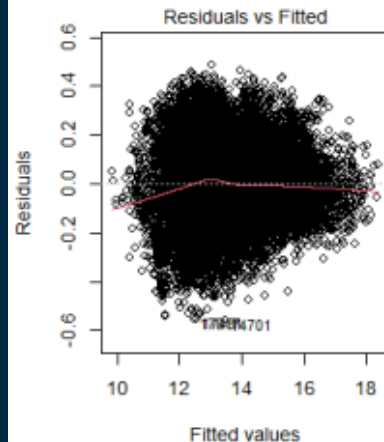
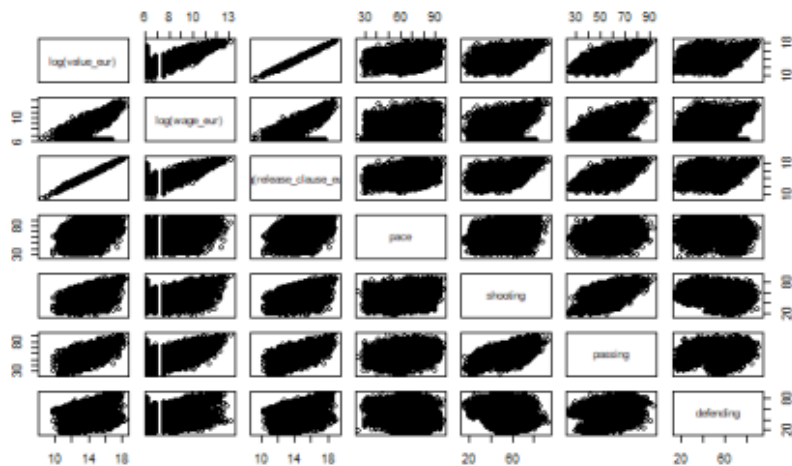
# Problems with the model

```
lm(value_eur ~ overall + wage_eur + release_clause_eur + pace + shooting +  
passing + dribbling + defending, data = Playas21)
```



# Final Model

$\text{lm}(\log(\text{value\_eur}) \sim \log(\text{wage\_eur}) + \log(\text{release\_clause\_eur}) + \text{pace} + \text{shooting} + \text{passing} + \text{defending}, \text{data} = \text{Playas21})$





# Summary Statistics

```
Call:
lm(formula = log(value_eur) ~ log(wage_eur) + log(release_clause_eur) +
    pace + shooting + passing + defending, data = Playas21)

Residuals:
    Min       1Q   Median       3Q      Max
-0.5601 -0.1169 -0.0114  0.1211  0.4865

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -0.1483387   0.0150693   -9.844  <2e-16 ***
log(wage_eur)  0.0556486   0.0014331   38.830  <2e-16 ***
log(release_clause_eur) 0.08904898 0.0016785  530.524  <2e-16 ***
pace           0.0014818  0.0001414   10.482  <2e-16 ***
shooting       0.0042496  0.0001864   22.794  <2e-16 ***
passing        0.0023768  0.0002342   10.147  <2e-16 ***
defending      0.0038047  0.0001328   28.643  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1683 on 15952 degrees of freedom
(2985 observations deleted due to missingness)
Multiple R-squared:  0.985,    Adjusted R-squared:  0.985
F-statistic: 1.744e+05 on 6 and 15952 DF,  p-value: < 2.2e-16
```

# Conclusion

- We reject the null hypothesis that the predictor variable is not correlated with the response



# Further Research

- Could do analysis on other datasets to see what attributes increases a players value over time.
- Attacking Crossing/ Skill Dribbling/ Movement Agility etc

