# Analyzing the Offensive Capabilities of Teams at the 2022 FIFA World Cup

Rahiq Raees

Danish Siddiqui

Sauhaard Walia

**Shaafay Alaman** 

**Acash Fernando** 

**Aaditya Samant** 

2024-04-07

# **Table of Contents**

#### **Abstract**

#### Introduction

#### Data

#### Methodology, Results, and Visualizations

- Expected Goals Model
- Expected Successful Passes Model
- Expected Assisted Goals Model

# **Summary and Conclusion**

References

#### **Appendix**

#### **Figures**

The table of contents is linked.

# **Abstract**

We apply several logistic regression models with the aim to measure the predicted offensive capabilities of teams at the 2022 FIFA World Cup against their actual offensive skill. The data used by our models is obtained from StatsBomb, a company that gathers event level data for a variety of sports. Library StatsBombR is installed in order to access the data. We answer our research prompt by analyzing expected goals (xG), expected successful passes (xP) as well as expected assisted goals (xAG) for each team in the tournament. The data contains a multitude of variables which help the regression models yield to the research prompt such as pass length, distance to goal, and pass angle. These variables are utilized to derive our xG, xP, and xAG models. Focussing on our xG model, we address our research prompt by comparing the expected goals for each team to their total goals actually scored. The teams whose expected goals are much lower than their total goals are said to overperform since they are doing better than we would expect, and the teams whose expected goals are much higher than their total goals are said to underperform since we would expect them to score more. The xP and xAG models' methodology is analogous. Omitted variable bias in our model is addressed by adding fixed team effects.

## Introduction

The FIFA World Cup is the largest sporting event in the world. FIFA claims that 1.5 billion people watched the final match live. The tournament takes place every four years, where the preceding three years are used as a qualification stage. 32 teams qualify and are eventually placed in knockout brackets to determine the winner. It is by far the most important football tournament due to its high viewership. Players that win/perform well in the World Cup tend to win the Ballon d'Or, a trophy awarded to the best player from the previous season, and it is considered by many to be the most prestigious individual award a player can obtain.

The World Cup is a very high stakes tournament due to its popularity and the nationalistic pride aspect, which makes it ripe for analysis. Answering our research prompt on how the predicted offensive capabilities of teams at our tournament compare to their actual offensive skill is important because it provides insight on strategic and performance analyses. For example, strategy wise, a manager will definitely want to replace inefficient forwards that have an expected goal tally that far exceeds their total goal tally. The same can be said for midfielders as their primary role is passing which is arguably the most important aspect of the sport. In terms of performance analysis, we can determine if expected successful passes, xG's and xAG's is enough to predict the winners, or does defense play a large enough role to combat this. Expanding on this, at the club level we see that forwards are paid substantially more than defenders, thus our analysis can determine if finances are being used wisely.

#### Data

All data was gathered using library(StatsBombR).

## [1] "Whilst we are keen to share data and facilitate research, we also urge you to be responsible with the data. Please credit StatsBomb as your data source when using the data and visit https://statsbomb.com/media-pack/ to obtain our logos for public use."

We used the FreeCompetitions() function from the library and then filtered out our desired competition.

```
matches <- FreeMatches(comp)</pre>
```

## [1] "Whilst we are keen to share data and facilitate research, we also urge you to be responsible with the data. Please credit StatsBomb as your data source when using the data and visit https://statsbomb.com/media-pack/ to obtain our logos for public use."

The FreeMatches() function extracts data for all 64 matches such as match id, home team id, away team id, home score, and away score.

```
main_data <- StatsBombFreeEvents(MatchesDF = matches, Parallel = T)</pre>
```

## [1] "Whilst we are keen to share data and facilitate research, we also urge you to be responsible with the data. Please credit StatsBomb as your data source when using the data and visit https://statsbomb.com/media-pack/ to obtain our logos for public use."

```
main_data = allclean(main_data)
```

The StatsBombFreeEvents() function extracts the event level data from all 64 games. This is the data that we use in our regressions. The allclean() function adds extra variables such as x/y coordinates.

The variables from our World Cup dataset are all mentioned and described in the following section.

# Methodology, Results, and Visualizations

## **Expected Goals Model**

We fist mention why expected goals is an important metric, because after all, the tournament has already taken place so we have data on total goals. The rebuttal to this is that expected goals enables us to determine which team is actually clinical in front of goal. For example, France was the top scorer in the tournament with 16 goals, but if their expected goals was 22, we would say France does not capitalize on their chances. We would say a team excels at finishing if their actual total goals is significantly more than their expected.

We now start our regression analysis.

The variables we think are important to include in our xG regression are

- shot.outcome.name: Dummy for goal/no goal
- shot.type.name: open play, free kick, or penalty
- · shot.body\_part.name: left foot, right foot, other
- location.x & location.y: (x,y) coordinate where shot was taken.
- DistToGoal & AngleToGoal
- DefendersInCone: Number of defenders within angle triangle.

Further explanations: In shot.body\_part.name, other refers mainly to headers, but can also include the very unlikely scenario of scoring from shoulder or torso. DistToGoal is computed by the euclidean distance formula and AngleToGoal is computed by an arctan() formula derived from basic trigonometry. DefendersInCone is an especially important varible as it captures the defensive pressure faced by a shooter. From where the shooters shoots, we imagine two line segments from his position connecting to the two goal posts. This is our angle triangle. The number of defenders within the angle triangle is the value that DefendersInCone takes on.

#### **Logistic Regression:**

Only 195 goals were scored in the tournament, which we think may be insufficient in training an accurate logistic regression model. 195 goals may not accurately represent all sorts of different angles, distances, and/or defensive pressures when a shot is attempted. We decided it is best to incorporate an independent training dataset to better depict our regression models. Thus, we pulled data from the 2015/2016 Premier League season (from Statsbomb) that has data on all 380 matches played, and consisted of 988 total goals scored during the season. Our training dataset contained the same variables as our dataset on the 2022 World Cup.

## [1] "Whilst we are keen to share data and facilitate research, we also urge you to be
responsible with the data. Please credit StatsBomb as your data source when using the da
ta and visit https://statsbomb.com/media-pack/ to obtain our logos for public use."
## [1] "Whilst we are keen to share data and facilitate research, we also urge you to be
responsible with the data. Please credit StatsBomb as your data source when using the da
ta and visit https://statsbomb.com/media-pack/ to obtain our logos for public use."

## [1] "Whilst we are keen to share data and facilitate research, we also urge you to be responsible with the data. Please credit StatsBomb as your data source when using the data and visit https://statsbomb.com/media-pack/ to obtain our logos for public use."

```
##
## Call:
  qlm(formula = goal ~ shot.type.name + shot.body part.name + location.x +
##
       location.y + DistToGoal + AngleToGoal + factor(DefendersInCone),
##
       family = binomial, data = temp data)
##
## Deviance Residuals:
      Min
                      Median
                                   30
                                           Max
##
                 10
## -1.8612 -0.4624 -0.3127 -0.1981
                                        3.1115
##
## Coefficients:
##
                                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)
                                  35.039377 196.983952
                                                         0.178 0.85882
## shot.type.nameFree Kick
                                 -18.529957 196.968769
                                                        -0.094
                                                                0.92505
## shot.type.nameOpen Play
                                 -20.394984 196.968671
                                                        -0.104
                                                                0.91753
## shot.type.namePenalty
                                 -18.127066 196.968985 -0.092 0.92667
                                                         8.341 < 2e-16 ***
## shot.body part.nameLeft Foot
                                   0.964914
                                              0.115683
## shot.body_part.nameOther
                                                         3.267 0.00109 **
                                   1.735834
                                              0.531321
## shot.body part.nameRight Foot
                                   1.051676
                                              0.105736
                                                         9.946 < 2e-16 ***
## location.x
                                  -0.120131
                                              0.016160 -7.434 1.06e-13 ***
## location.y
                                  -0.014011
                                              0.011351 - 1.234 0.21710
## DistToGoal
                                              0.016454 - 15.694 < 2e-16 ***
                                  -0.258221
## AngleToGoal
                                   0.001699
                                              0.002506
                                                         0.678 0.49781
## factor(DefendersInCone)1
                                  -0.627614
                                              0.085361 -7.352 1.95e-13 ***
## factor(DefendersInCone)2
                                  -0.867202
                                              0.123357 -7.030 2.07e-12 ***
## factor(DefendersInCone)3
                                  -1.402435
                                              0.217258 -6.455 1.08e-10 ***
## factor(DefendersInCone)4
                                                        -4.495 6.95e-06 ***
                                  -1.199752
                                              0.266904
## factor(DefendersInCone)5
                                                        -3.975 7.04e-05 ***
                                  -1.823940
                                              0.458852
## factor(DefendersInCone)6
                                  -1.774324
                                              0.764556 - 2.321 \ 0.02030 *
## ---
                   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Signif. codes:
##
## (Dispersion parameter for binomial family taken to be 1)
##
##
                                       degrees of freedom
      Null deviance: 6423.2
                              on 9898
## Residual deviance: 5276.4
                              on 9882
                                       degrees of freedom
## AIC: 5310.4
##
## Number of Fisher Scoring iterations: 10
```

As expected, all factors of DefendersInCone, location.x, and DistToGoal is significant. Of course defensive pressure will have a major impact on goal scoring and of course the farther you are from the net, the harder it is to score.

shot.body.part.name is also significant. We attribute this to the fact that is it much harder to score with your non-dominant foot and to the effectiveness of headers taken near the goal.

We are surprised to find that location.y and AngleToGoal are not significant. A plausible explanation is that enough shots not being attempted at extreme angles and y- values.

shot.type.name is also an insignificant variable. We expect Penalty and Free Kick to surely be significant since the chance of scoring is quite high for penalty shots and quite low for free kicks. An explanation for this is that the DistToGoal and DefendersInCone variables may be overriding the effect of shot.type.name. Taking penalties as an example, the small distance to goal and no defenders in cone may be leading us to assign very high probabilities to these shots, thus we are miss-classifying the missed penalties, hence the predictor does not predict well.

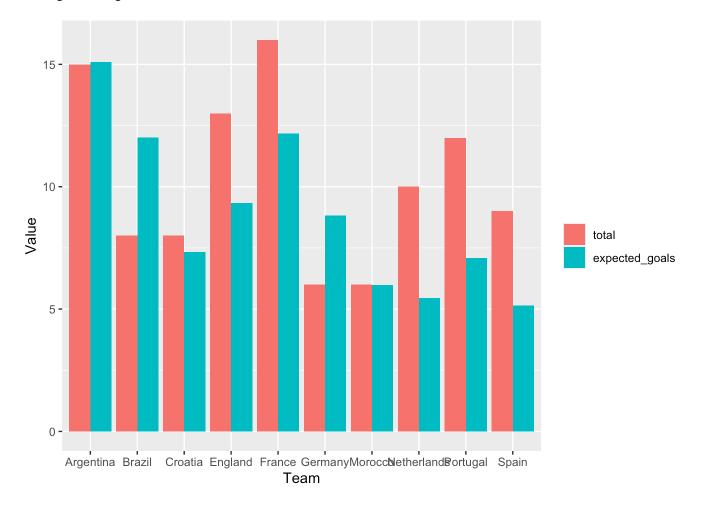
```
\label{eq:my_logit_2} $$ my_logit_2 <- glm(data= temp_data,goal \sim shot.type.name + shot.body_part.name + location. $$ x + location.y + DistToGoal + AngleToGoal + factor(DefendersInCone) + factor(team.name), family=binomial) $$
```

Omitted variable bias is addressed by adding team fixed effects.

```
##
                      (Intercept)
                                         shot.type.nameFree Kick
                     35.039376836
                                                    -18.529956703
##
##
         shot.type.nameOpen Play
                                           shot.type.namePenalty
##
                    -20.394984098
                                                    -18.127066209
    shot.body_part.nameLeft Foot
##
                                        shot.body part.nameOther
##
                      0.964914452
                                                      1.735833546
##
   shot.body_part.nameRight Foot
                                                       location.x
##
                      1.051675862
                                                    -0.120130992
##
                       location.y
                                                       DistToGoal
##
                     -0.014010594
                                                    -0.258220815
##
                      AngleToGoal
                                        factor(DefendersInCone)1
                      0.001698893
                                                     -0.627613529
##
##
        factor(DefendersInCone)2
                                        factor(DefendersInCone)3
##
                     -0.867201596
                                                     -1.402434846
##
        factor(DefendersInCone)4
                                        factor(DefendersInCone)5
##
                     -1.199751634
                                                    -1.823939690
        factor(DefendersInCone)6
##
                     -1.774324052
##
```

```
(Intercept)
                                         shot.type.nameFree Kick
##
##
                     35.032175532
                                                    -18.613247828
##
         shot.type.nameOpen Play
                                           shot.type.namePenalty
                    -20.454883103
##
                                                    -18.161499618
##
    shot.body_part.nameLeft Foot
                                        shot.body part.nameOther
##
                      0.946886104
                                                      1.677960572
##
   shot.body part.nameRight Foot
                                                       location.x
                      1.043632751
                                                     -0.119932477
##
                                                       DistToGoal
##
                       location.y
##
                     -0.013913935
                                                     -0.258444246
                                        factor(DefendersInCone)1
##
                      AngleToGoal
                      0.001698364
                                                     -0.626101868
##
        factor(DefendersInCone)2
                                        factor(DefendersInCone)3
##
##
                     -0.863509407
                                                     -1.373425744
        factor(DefendersInCone)4
                                        factor(DefendersInCone)5
##
                     -1.155504526
                                                    -1.791434380
##
##
        factor(DefendersInCone)6
##
                     -1.734492999
```

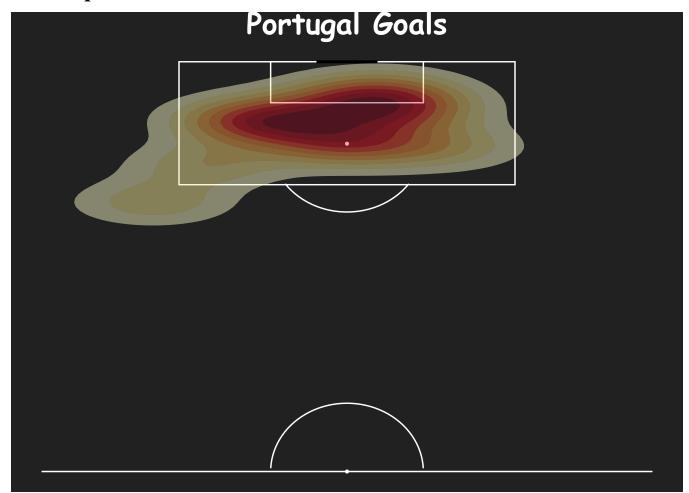
We see that the coefficients in both models are very similar to each other. Since fixed effect models allow us to control for variables that are constant over time, it seems that these variables are not important for the estimates we're generating.



Portugal and Netherlands overperformed by the most. They scored a much higher number of goals than our model predicted. See: angle and distance of Portugal's goals. These are goals that are NOT expected to go in.

Brazil and Canada were the largest underperformers. This could be due to the fact that they created a lot of opportunities, but failed to capitalize.

#### **Heatmaps**:



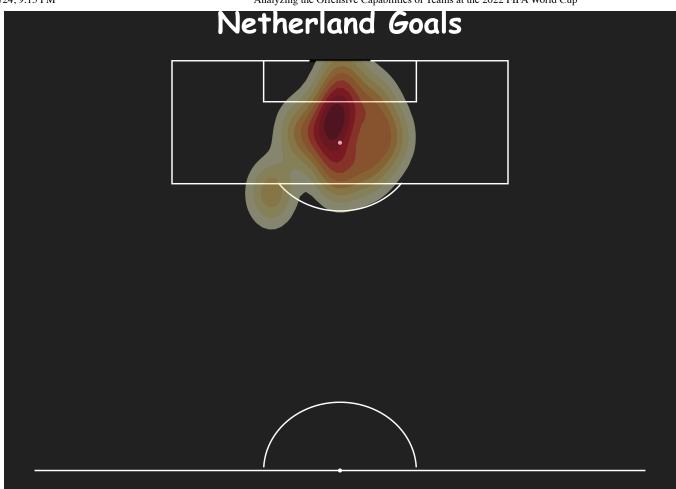
Top Overpeformer:

Portugal -

Total Goals = 12

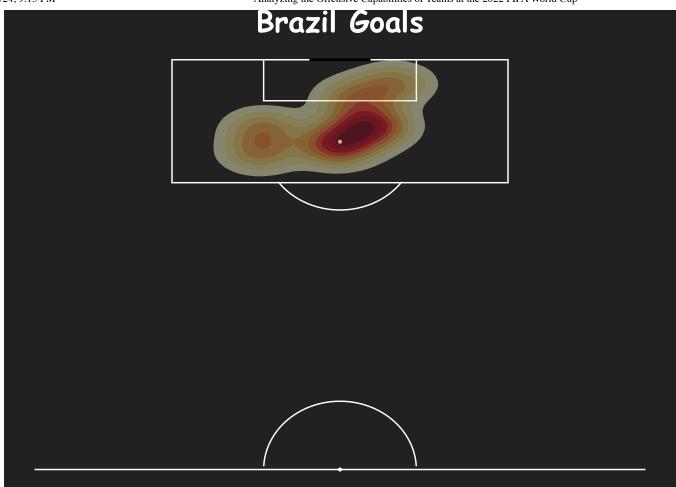
xG = 7.1

Overperformed by 4.9 goals



Netherlands -Total Goals = 10 xG = 5.46

Overperformed by 4.54 goals



Top Underperformer:

Brazil -

Total Goals = 8

xG = 12.01

Underperformed by 4.01 goals

Brazil's underperformance can be chalked down to them being absolute favourites and injuries. They were loaded with elite strikers such as Vini & Neymar Jr. Not only did they fail to capitalize on chances, but faced a pivotal injury.

Neymar missed 2 out of Brazil's 5 games in the WC. He is currently Brazil's all time top-scorer.

#### **Expected Successful Passes Model**

The orchestration of the offense is ideally controlled by your midfielders. The entire footballing world is seeing a revolution towards Johan Cruyff's style of play. The tactical mastermind's style hinged on successful passing. In this day and age, the teams that pass more tend to have the best offenses.

Just take a look at the likes of Man City, Barcelona, Arsenal.

We built our xP model using world cup data, as the number of passes were enough to create an accurate model - unlike xG.

We now start our regression analysis.

The variables we think are important to include in our xP regression are

- pass.length: the length of a pass computed via Euclidean distance formula.
- pass.angle: the angle at which a pass is being given
- · pass.height.name: the height of the ball when passed

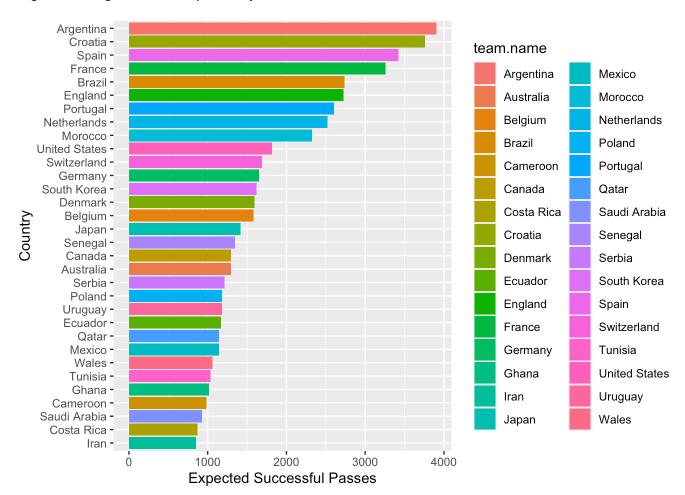
#### **Logistic Regression**:

## [1] "Whilst we are keen to share data and facilitate research, we also urge you to be responsible with the data. Please credit StatsBomb as your data source when using the data and visit https://statsbomb.com/media-pack/ to obtain our logos for public use."

```
##
## Call:
  glm(formula = successful_pass ~ pass.length + pass.angle + pass.height.name,
##
       family = binomial, data = passes)
##
## Deviance Residuals:
                      Median
                                   30
##
       Min
                 10
                                           Max
## -2.3135
             0.3799
                      0.3831
                               0.3865
                                        1,2916
##
## Coefficients:
##
                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)
                              2.5796239 0.0223786 115.272
                                                              <2e-16 ***
## pass.length
                             -0.0003553 0.0007877 -0.451
                                                               0.652
                                                               0.299
## pass.angle
                              0.0081346 0.0078380
                                                      1.038
## pass.height.nameHigh Pass -2.7998496 0.0300658 -93.124
                                                              <2e-16 ***
## pass.height.nameLow Pass -1.7017281 0.0312923 -54.382
                                                              <2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
       Null deviance: 64096
##
                             on 68514
                                       degrees of freedom
## Residual deviance: 50167 on 68510
                                       degrees of freedom
## AIC: 50177
## Number of Fisher Scoring iterations: 5
```

We see that both factors of pass.height.name are significant predictors. This makes sense since low passes are usually easier to control than airborne passes.

We are surprised to see that pass.length and pass.angle are insignificant. This may be due to the fact that a substantial majority of our data is comprised of passes with similar lengths and angles, and so the extreme lengths and angles are not captured by our model well.



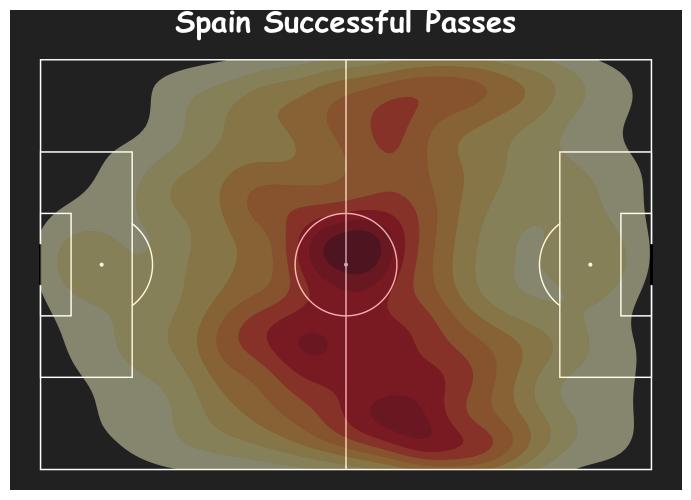
As mentioned, midfielders are important! Hence, the reason why Argentina and Croatia have the highest successful expected passes (elite players like Messi and Modric).

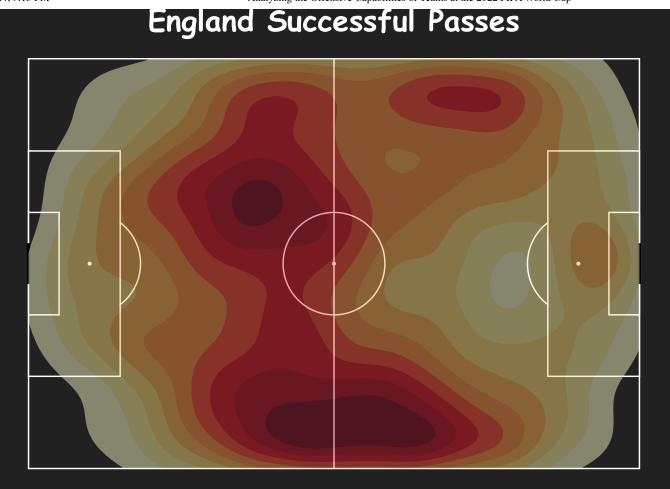
This midfielders deliver passes so beautifully that the angle, length, height etc. is perfect.

team.name	actual_successful_passes	expected_successful_passes	difference
Spain	3540	3421.153	118.84704
England	2800	2724.032	75.96780
Portugal	2655	2602.614	52.38610
Croatia	3796	3758.172	37.82773
Brazil	2776	2738.889	37.11062
Argentina	3939	3902.774	36.22565
team name	actual successful nasses	expected successful passes	difference

team.name	actual_successful_passes	expected_successful_passes	difference
Costa Rica	830	869.9825	-39.98247
Tunisia	996	1037.3625	-41.36254
Saudi Arabia	880	929.6270	-49.62704
Iran	792	850.3625	-58.36252
Japan	1355	1415.2460	-60.24601

### **Heatmaps**:





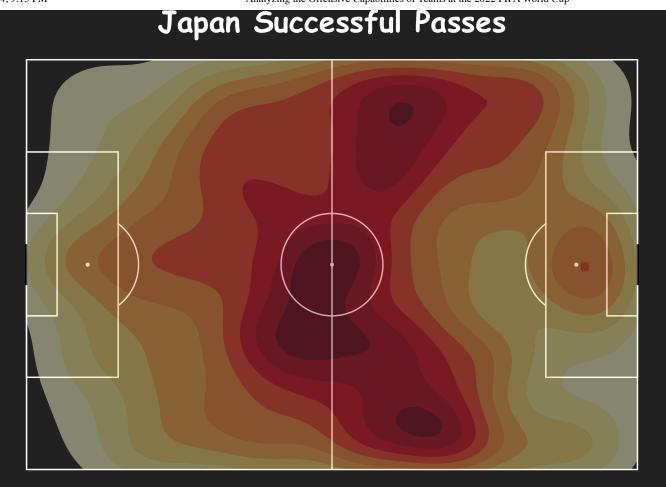
Spain has over performed the most when it comes to successful passing. They achieved 119 passes more than what was expected of them.

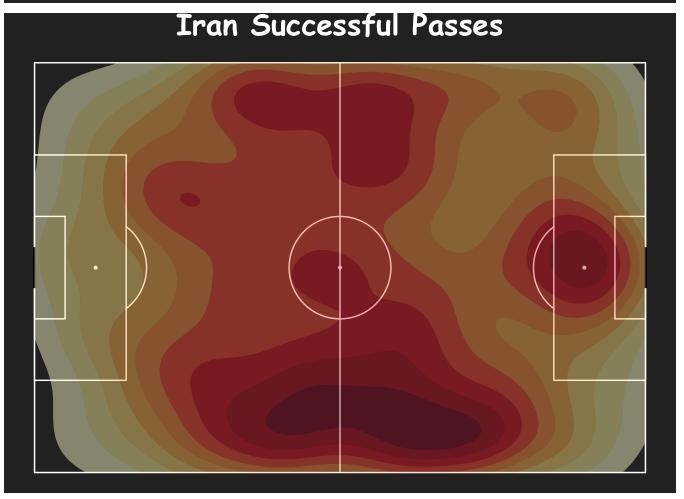
As seen in the heatmap, their passing is heavily concentrated in the midfield. This is also contributed by strikers on the left who drop back to receive passes. Despite Spain having the high benchmark of third highest xP, they still managed to achieve the highest overperformance.

The second highest over performer is England. They have several hotspots of passing.

This makes sense as the Trent can be seen receiving the ball high up the pitch on the left side.

England could have also overperformed due to their strong midfield passers such as Bellingham, Rice, and Mount.





Japan was the biggest passing underperformer at the World Cup. This means their goal-scoring carried them in the world cup - gaining some upsets such as their win against Spain and Germany.

Iran underperformed the most after Japan. This can be due to a number of reasons. Iran tends to underperform on the big stage despite their potential. It tends to happen at several competitions despite being ranked 20th in the world.

#### **Expected Assited Goals Model**

xAG quantifies the probability of an assist resulting in an expected goal, rather than just a scoring chance.

It investigates the contribution of an assist towards the likelihood of a goal being scored.

Through this model we gain a deeper understanding on the impact that each assist has in the game.

The same Premier League training data was used as our xG model to train the model.

We now start our regression analysis.

The variables we think are important to include in our xP regression are

- pass.length: the length of a pass computed via Euclidean distance formula.
- pass.angle: the angle at which a pass is being given
- · pass.height.name: the height of the ball when passed

#### **Logistic Regression:**

A beta regression was used for this model using the following coefficients:

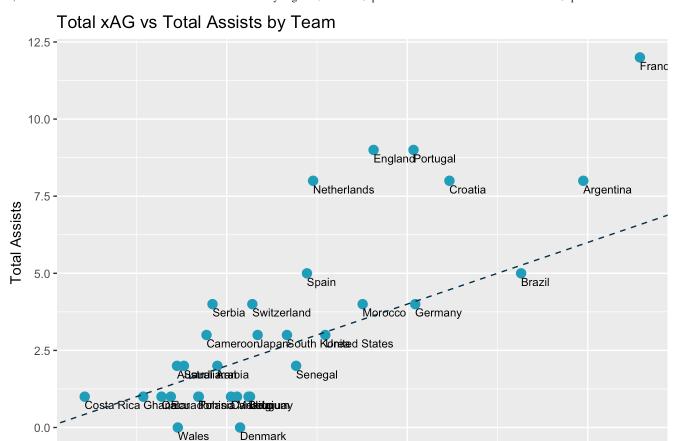
- shot.statsbomb\_xg: xG that is calculated by StatsBomb using advanced neural networks.
- pass.length: the length of a pass computed via Euclidean distance formula.
- · pass.height.name: the height of the ball when passed
- · pass.technique.name: type of pass

```
##
## Call:
## betareg(formula = shot.statsbomb xg \sim pass.length.x + pass.height.name.x +
      pass.technique.name.x, data = assist events)
##
##
## Standardized weighted residuals 2:
##
      Min
               1Q Median
                              30
                                    Max
## -4.9063 -0.4908 -0.0485
                          0.4481
                                 4.8561
##
## Coefficients (mean model with logit link):
##
                                    Estimate Std. Error z value Pr(>|z|)
                                  -2.3386549 0.0577128 -40.522 < 2e-16 ***
## (Intercept)
## pass.length.x
                                  ## pass.height.name.xHigh Pass
                                   0.3944904 0.0267537 14.745 < 2e-16 ***
## pass.height.name.xLow Pass
                                   0.2078877 0.0297428
                                                       6.990 2.76e-12 ***
## pass.technique.name.xOther
                                  -0.2150756 0.0518437 -4.149 3.35e-05 ***
## pass.technique.name.xOutswinging
                                  ## pass.technique.name.xStraight
                                  -0.1401511 0.0910412 -1.539
                                                                 0.124
## pass.technique.name.xThrough Ball 0.6991345 0.0635983 10.993 < 2e-16 ***
##
## Phi coefficients (precision model with identity link):
##
        Estimate Std. Error z value Pr(>|z|)
## (phi)
         16.1147
                    0.2967
                             54.31
                                    <2e-16 ***
## ---
                 0 '*** 0.001 '** 0.01 '* 0.05 '. 0.1 ' 1
## Signif. codes:
##
## Type of estimator: ML (maximum likelihood)
## Log-likelihood: 1.025e+04 on 9 Df
## Pseudo R-squared: 0.1103
## Number of iterations: 17 (BFGS) + 2 (Fisher scoring)
```

Besides pass.technique.name.xStraight, we see that all our predictors are significant. pass.technique.name.xStraight has a 12% p-value which is not too small, so it may still be somewhat significant. The rest of the predictors have very small p-values, hence they are for sure significant.

4

Total xAG

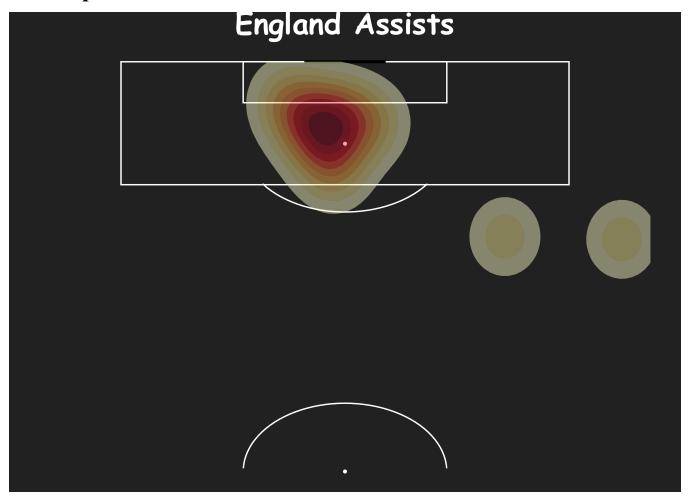


As can be seen in the chart on the right, the top two overperformers were France and England. The most underperforming teams when it came to this metric were Denmark and Uruguay.

2

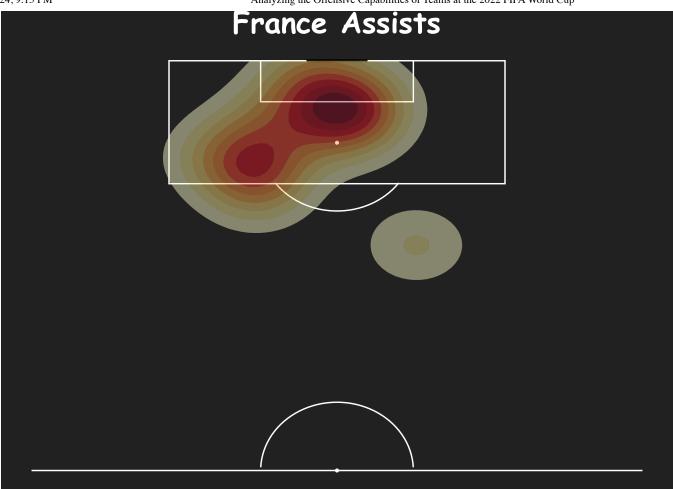
6

#### **Heatmaps**:



England's assists are uniquely far away from the goalscorer. Observe the two heat spots outside the penalty area. The model would have predicted these assists to diminish the outcome of a goal.

This gives us some extra context as to why England outperformed



France's actual assists were concentrated down the left hand side (guess who?).

Given pass length and the technique required from these locations (most likely outswinging, or inswinging), the negative coefficients may have driven the aggregate xAG down, resulting in overperformance.

# **Summary and Conclusion**

At the 2022 World Cup, Argentina were top. The team performed amazingly, with the "little boy from Rosario, Argentina," being the hero of the tournament. They etched their names into footballing history, proving to everyone watching - miracles do happen. However, they performed in line with what our models would expect, as reflected by the slim residuals.

Overall, Portugal, Netherlands, France, England, and Spain had the largest residuals with our regressions this world cup. These teams played brilliantly. In terms of overperformance, we can conclude there was one team ahead of the rest in terms of offensive prowess: the Three Lions. England rose ahead of their expectations, as demonstrated by their large residuals from our xAG and xP models. One may think this is because our regressions predicted a low baseline for England, making it easy to overperform relative to expectation - but this was not the case. The xAG overperformance reflects good finishing in difficult positions by clinical attackers - such as Harry Kane - scoring from assists in positions from assists where they are unlikely to get assists. On the other hand, outperforming xP was down to their terrific midfield. If they did not play the breathtaking French team in the guarter finals, who knows how much further they could have made it.

In fact, this could be something worth modeling. Our research could have been developed further into the "what if?" question. Using predictive regressions, we could have taken note of the overperformers of the tournament. After which, instead of ending it there, we could have investigated - if the teams kept going (i.e. not losing, consequently getting kicked out) further into the tournament - would they have defeated their opponents? While this does not factor in injuries, field influence, crowd, and arbitrary variables such as weather - it would have been interesting to see the results. It would have painted the significance of delivering during clutch key moments (seen time and time again by Argentina this tournament). Those could be the difference between a World Cup trophy, and being forgotten. We could have answered the question: "Would have overperformers such as England or Portugal taken down teams further down the road, had they not lost earlier in the tournament?"

Lastly, for each regression we noted that some p-values were quite surprising. Predictors that we would expect to be significant turned out not to be. We believe that if we had more data on the tournament, then these predictors would have been significant.

#### References

StatsBomb. (2019). Using StatsBomb Data In R. Retrieved from https://statsbomb.com/wp-content/uploads/2019/12/Using-StatsBomb-Data-In-R.pdf (https://statsbomb.com/wp-content/uploads/2019/12/Using-StatsBomb-Data-In-R.pdf)

FIFA.com. (2022, December 20). One month on: 5 billion engaged with the FIFA World Cup Qatar 2022™. FIFA. Retrieved from https://inside.fifa.com/tournaments/mens/worldcup/qatar2022/news/one-month-on-5-billion-engaged-with-the-fifa-world-cup-qatar-2022-tm

(https://inside.fifa.com/tournaments/mens/worldcup/qatar2022/news/one-month-on-5-billion-engaged-with-the-fifa-world-cup-qatar-2022-tm)

NPR. (2022, December 18). World Cup final pits Argentina against France, Messi against Mbappe. NPR. Retrieved from https://www.npr.org/2022/12/18/1143933269/world-cup-final-argentina-france-messi-mbappe (https://www.npr.org/2022/12/18/1143933269/world-cup-final-argentina-france-messi-mbappe)

# **Appendix**

#### Overperformers and underperformers in terms of expected goals.

team.name <chr></chr>	<b>total</b> <int></int>	expected_goals <dbl></dbl>	total - expected <dbl></dbl>
1 Portugal	12	7.093666	4.906334
2 Netherlands	10	5.462007	4.537993
3 Spain	9	5.143043	3.856957
4 France	16	12.171358	3.828642
5 England	13	9.338299	3.661701
6 Croatia	8	7.327844	0.672156
6 rows			

	team.name <chr></chr>	total <int></int>	expected_goals <dbl></dbl>	total - expected <dbl></dbl>
5	England	13	9.338299	3.661701438
6	Croatia	8	7.327844	0.672156050
7	Morocco	6	5.993917	0.006083302
8	Argentina	15	15.087809	-0.087809297
9	Germany	6	8.818260	-2.818260290
10	Brazil	8	12.007859	-4.007859149
6 rov	WS			

#### Overperformers and underperformers in terms of expected assisted goals.

team.name	total_xAG	total_assists
France	6.577922	12
Argentina	5.950341	8
Brazil	5.261496	5
Croatia	4.467386	8
Germany	4.088254	4
Portugal	4.069394	9

team.name	total_xAG	total_assists
Wales	1.4576353	0
Australia	1.4490505	2
Ecuador	1.3803797	1
Qatar	1.2773385	1
Ghana	1.0751753	1
Costa Rica	0.4264636	1

# **Figures**

