An Analysis of the 2018/2019 Fifa World Cups

Billy Hanan
Student No. : 18179797
Postgrad Diploma in Data Analytics
Advanced CRM: CA 2
07/12/2019

England	FA Women's Super League
International	FIFA World Cup
Spain	La Liga
United States of America	NWSL
International	Women's World Cup

StatsBomb Event Data

timestamp	location	type.name	team.name	player.name	pass.recipient.name	pass.outcome.name	pass.length	pass.end_location	shot.outcome.name
<chr>	<list>	<chr>	<chr>	<chr>	<chr>	<chr>	<dbl>	<list>	<chr>
00:00:56.190	64.3, 61.2	Pass	Korea Republic Women's	Bo-Ram Hwang	Seol-Bin Jung	Incomplete	45.37632	99.2, 32.2	NA
00:00:58.771	97.3, 34.6	Ball Receipt*	Korea Republic Women's	Seol-Bin Jung	NA	NA	NA	NULL	NA
00:00:58.771	21.8, 48.8	Clearance	France Women's	Griedge Mbock Bathy Nka	NA	NA	NA	NULL	NA
00:01:00.181	83.4, 39.3	Ball Recovery	Korea Republic Women's	So-Hyun Cho	NA	NA	NA	NULL	NA
00:01:00.221	84.8, 42.1	Shot	Korea Republic Women's	So-Hyun Cho	NA	NA	NA	NULL	Blocked

Match Feature Generation

'match_id' 'match_date' 'team.name' 'score' 'Shots' 'ShotsOnTarget' 'AvgShotDistance' 'XG' 'Possession' 'PctPossession'

'avg.pass.length' 'avg.pass.length.z1' 'avg.pass.length.z2' 'avg.pass.length.z3' 'length.of.resultant.pass' 'dirn.of.resultant.pass'

'length.of.resultant.pass.z1' 'dirn.of.resultant.pass.z1' 'length.of.resultant.pass.z2' 'dirn.of.resultant.pass.z2' 'length.of.resultant.pass.z3'

'dirn.of.resultant.pass.z3' 'z1.passes' 'z2.passes' 'z3.passes' 'z1.pass.pct' 'z2.pass.pct' 'z3.pass.pct' 'Result' 'Opp_team.name' 'Opp_score'

'Opp_Shots' 'Opp_ShotsOnTarget' 'Opp_AvgShotDistance' 'Opp_XG' 'Opp_Possession' 'Opp_PctPossession' 'Opp_avg.pass.length'

'Opp_avg.pass.length.z1' 'Opp_avg.pass.length.z2' 'Opp_avg.pass.length.z3' 'Opp_length.of.resultant.pass' 'Opp_dirn.of.resultant.pass'

'Opp_length.of.resultant.pass.z1' 'Opp_dirn.of.resultant.pass.z1' 'Opp_length.of.resultant.pass.z2' 'Opp_dirn.of.resultant.pass.z2'

'Opp_length.of.resultant.pass.z3' 'Opp_dirn.of.resultant.pass.z3' 'Opp_z1.passes' 'Opp_z2.passes' 'Opp_z3.passes' 'Opp_z1.pass.pct'

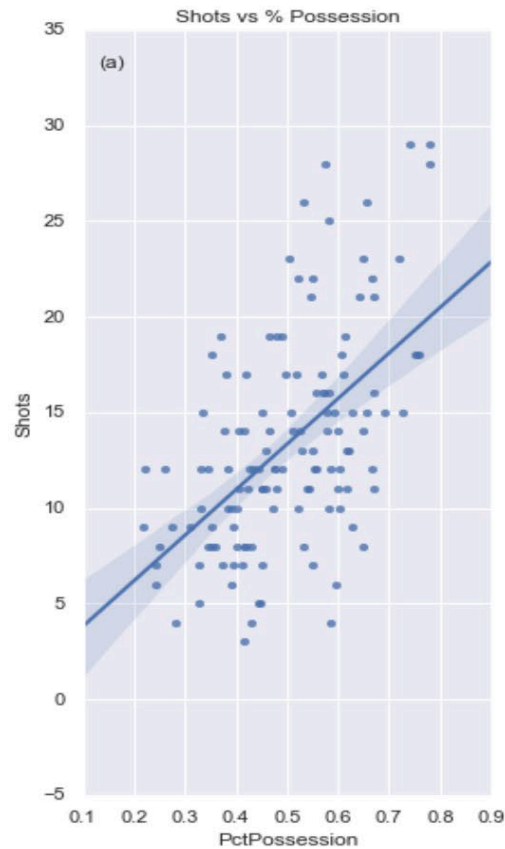'Opp_z2.pass.pct' 'Opp_z3.pass.pct' 'Opp_Result'

Match Record

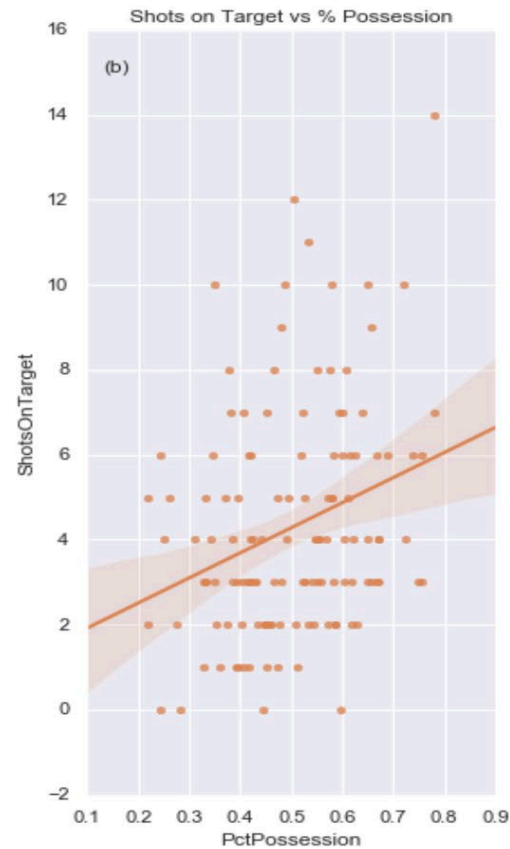|	match_date	team.name	score	Shots	PctPossession	Opp_team.name	Opp_score	Opp_Shots	Opp_PctPossession	y
62	2018-07-14	Belgium	2	11	0.421453	England	0	15	0.578547	1
63	2018-07-15	France	4	8	0.345154	Croatia	2	15	0.654846	1

DataSets:

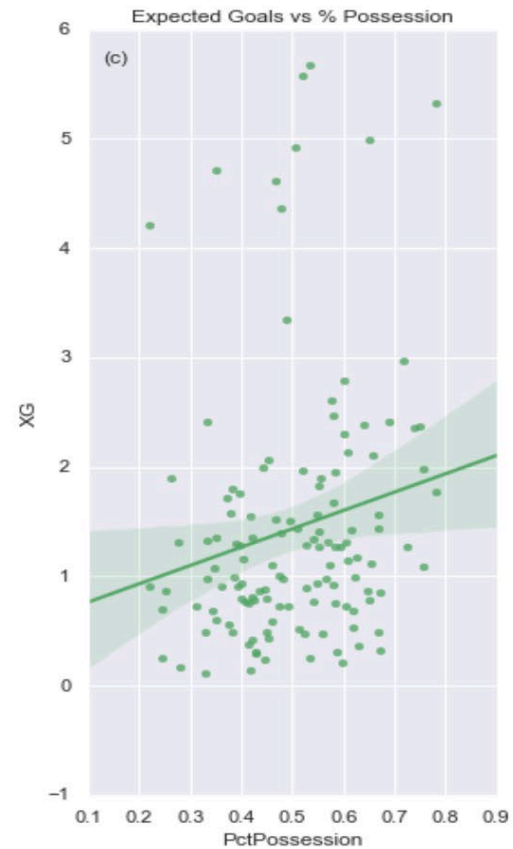- World Cup 2018: 64 Match Records
- World Cup 2019: 52 Match Records

WC 2018: Goals Scored vs Ball Possession
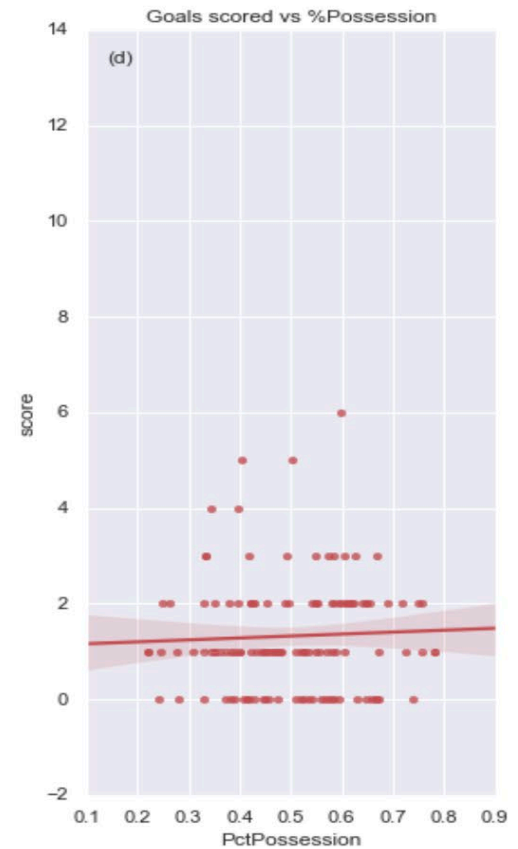

m=23.81
p=1.69e-11

m=5.90
p=0.00091

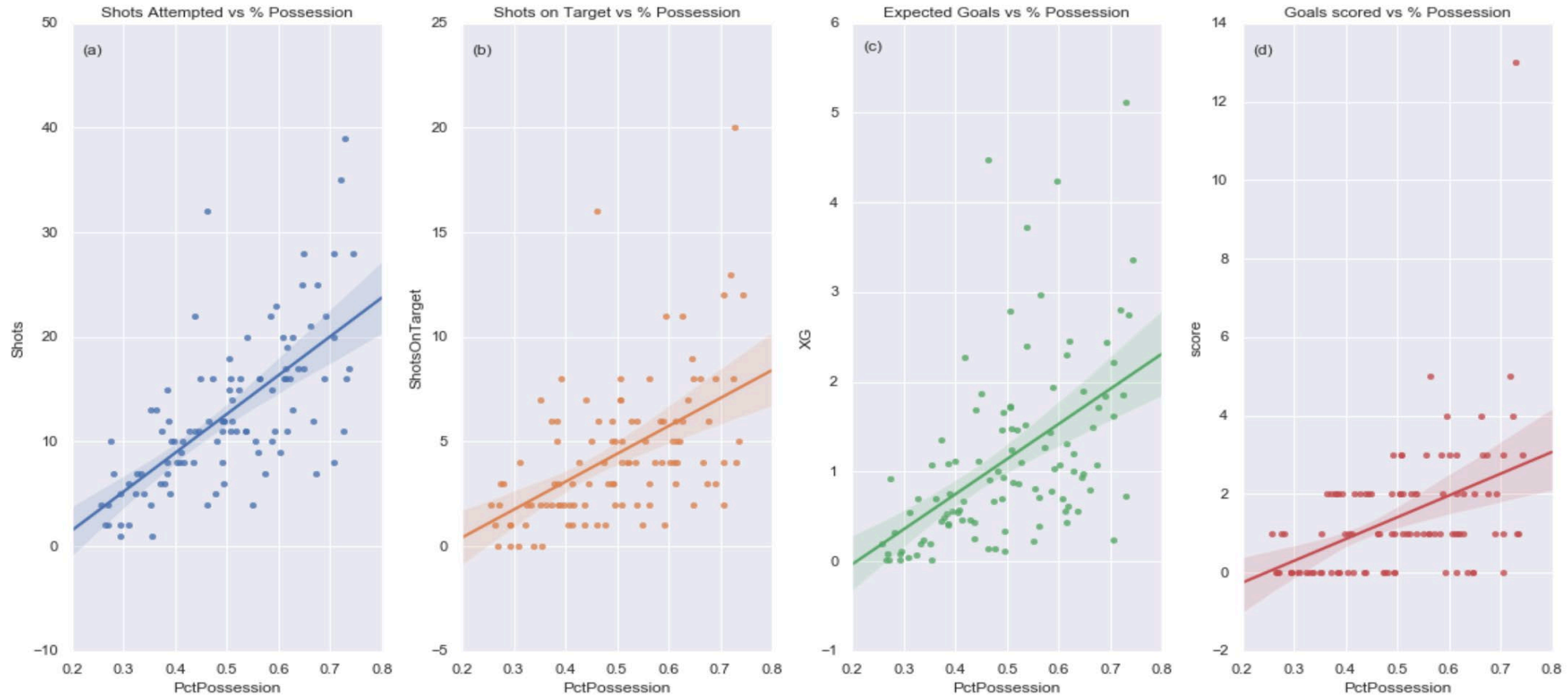m=1.66
p=0.03427

m=0.41
p=0.607

WC 2019: Goals Scored vs Ball Possession

m=37.08
p=4.25e-15

m=13.32
p=1.20e-08

m=3.91
p=1.58e-08

m=4.31
p=4.6e-7

WC 2018 : Logistic Regression Model

```
------------------------------------------------------------------------------
                              Coef.   Std.Err.      z      P>|z|    [0.025   0.975]
------------------------------------------------------------------------------
Opp_ShotsOnTarget            -0.6368    0.1957  -3.2541  0.0011  -1.0204  -0.2533
Opp_z3.pass.pct              12.8069    4.0524   3.1603  0.0016   4.8644  20.7495
length.of.resultant.pass.z3  -0.4528    0.2175  -2.0823  0.0373  -0.8791  -0.0266
==============================================================================
```

Prediction Accuracy = 0.78125

WC 2019 : Logistic Regression Model

```
---------------------------------------------------------------------
                    Coef.   Std.Err.     z      P>|z|    [0.025   0.975]
---------------------------------------------------------------------
z3.passes          0.0233    0.0071   3.3026  0.0010   0.0095   0.0371
Opp_ShotsOnTarget -0.6091    0.1813  -3.3586  0.0008  -0.9645  -0.2536
=====================================================================
```

Prediction Accuracy = 0.884615384615

Conclusion

- Ball Possession was an important factor in the women's World Cup 2019
- A tight defence is the single most important characteristic for a winning football team