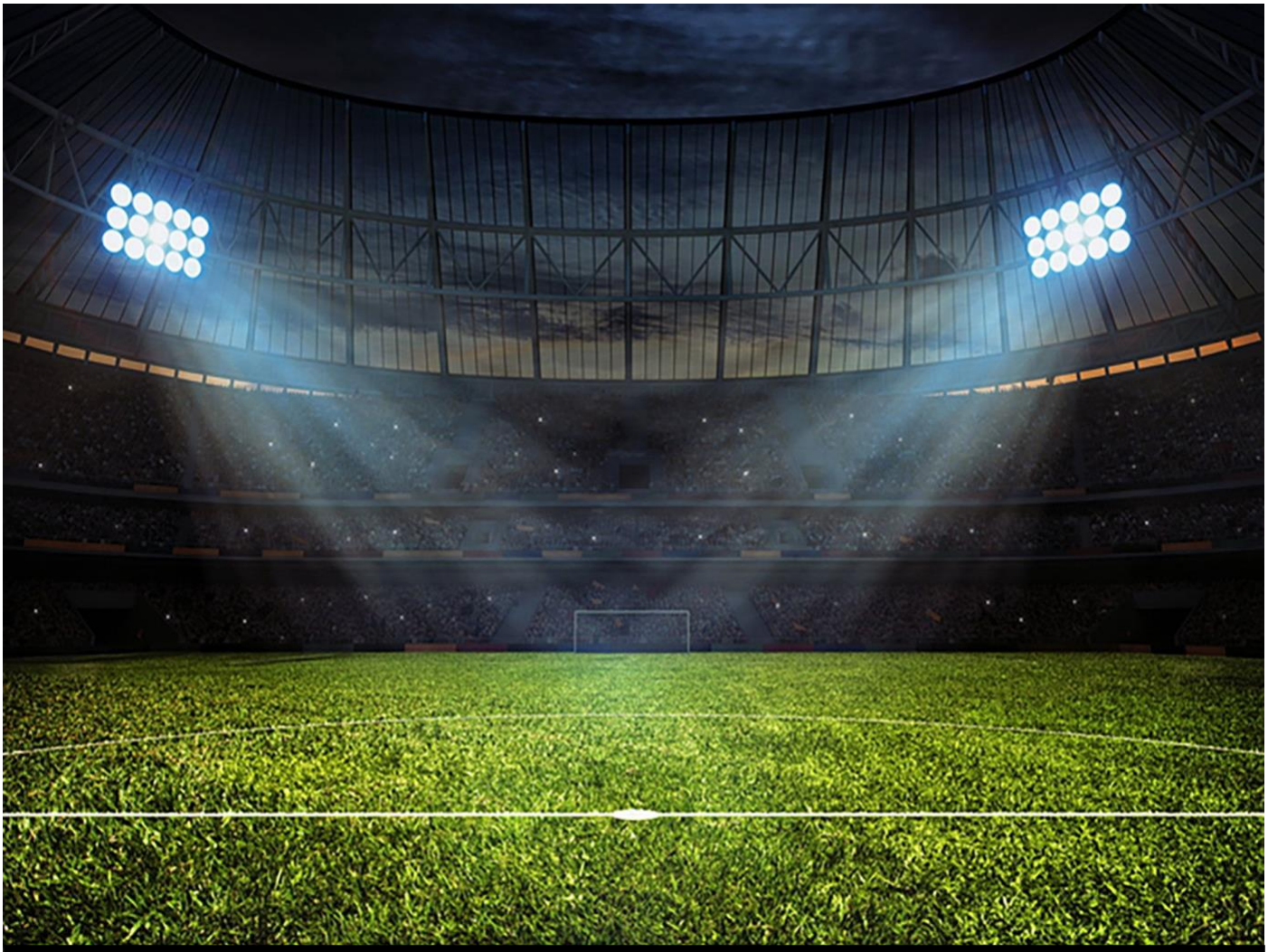


Executive Report for Rarita's National Football Team Plan

2022 Student Research Case Study Challenge
Football/ Soccer

MARBS CONSULTING

Team: MARBS Consulting (UNSW)

Margarita Psaras, Athena Yuan, Raymond Li, Isabella Katsiris, Skyler Zou

TABLE OF CONTENTS

01 OBJECTIVES OF ANALYSIS

MARBS Consulting has been hired to form a competitive international football team for Rarita and to analyse the impact of building a football "brand" on the economy. The long-term objective is to add positive economic impact through such branding from competitiveness, higher tourism rates, and creating new jobs.

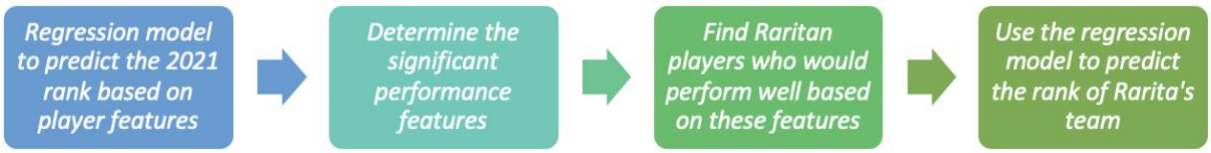
The metrics of the team's competitiveness will be measured as the probability of ranking top ten within five years and a high probability of obtaining first within ten years. The objective of building a team is to maximise these probabilities of success, whilst remaining within the budget of 995,000,000 Doubloons for ten years.

Additionally, the impact of building Rarita's Football "brand" will be done through exploration of revenues and expenses. Quantitatively, key economic indices such as GDP and probability distributions will provide a measure of the positive economic impact. The qualitative analysis will be done through risk management conventions such as the Risk Categorisation and Definition (RCD) tool and SWOT analyses.

02 TEAM SELECTION

The team selection was largely based on the player data which provided performance statistics of players and characteristics such as age, country, and salary. A regression model was used to accurately predict the 2021 rank of nations in tournament and then determine the statistically significant performance features that contributed to top performing nations. From there, Rarita's top team players were selected, and a prediction was output.

Figure 2.1: Summary of selection process



CRITERIA FOR SELECTION

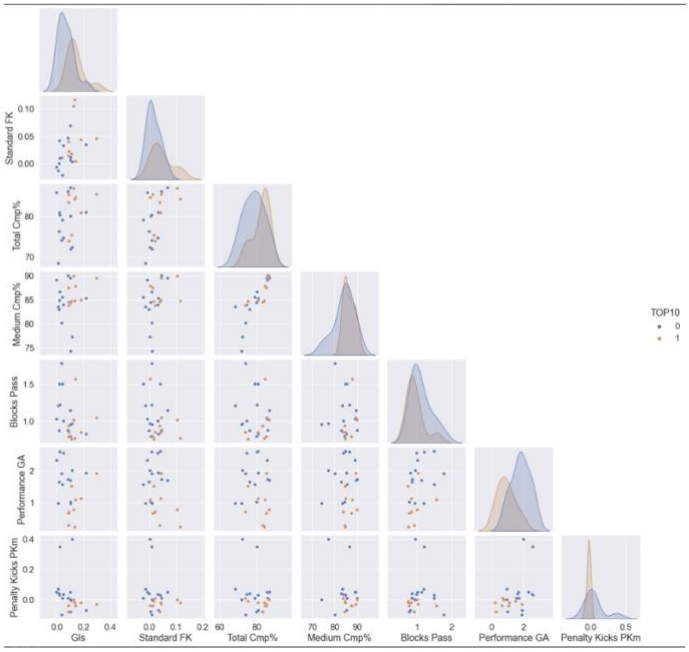
A lasso regression model using performance features (e.g., tournament shooting, passing etc.) of each nation was used to predict their 2021 team rank. As the lasso regression model penalises non-significant features, we were able to obtain the most important variables based on their contribution to predicting 2021 tournament rank. The most important selected variables were:

- Goals
- Shots from free kicks
- Percentage of passes completed
- Percentage of passes completed within 15-30 yards
- Number of times where a pass was blocked
- Goals scored against
- Penalty kicks missed

(See appendix B for key results.)

We analysed how these predictor variables differentiated the top nations in the top 10 performing teams to the rest. For example, in Figure 2.2 the distribution of goals scored on average is higher for those in top 10.

Figure 2.2: The distribution of player attributes in the pairwise plots



Once these predictors were confirmed to influence the predictability of being in top 10, we used the provided league data to select the ideal players. We also added qualitative criteria (Table 2.1) to meet our long-term objectives and a salary variable to ensure that it was within the budget provided. Below are the top ranked Raritan players (Table 2.2).

Table 2.1 Qualitative criteria and reasoning to filter players

Criteria	Reason
Age must be below 30	Average age of retirement is 35 so to build a long-term team
Nation must be Rarita	Building the Football "brand" in Rarita is more effective if Raritan's are the players selected.

Table 2.2 Ranked Raritan players based on selected features

	Player	Pos	Gls	Standard FK	Total Cmp%	Medium Cmp%	Blocks Pass	Performance GA	Penalty Kicks PKm	Value
0	K. Kazlo?	FW	4.832101	3.937875	-0.578374	-0.882552	-0.945175	NaN	NaN	-11.336899
1	B. Ampofo	GK	-0.118982	0.540325	0.321024	1.613995	-1.150298	-1.147161	-0.909313	-8.482538
4	J. Namirembe	MF	2.115043	2.117759	-0.915648	-0.898199	-0.375389	NaN	NaN	-4.578821
5	H. Mubaiwa	FW	2.598076	0.054960	-0.253793	-0.867884	-1.013549	NaN	NaN	-4.269473
9	G. Jankowski	MF	2.537697	-0.309063	0.062628	-0.191186	0.114628	NaN	NaN	-3.533854
11	P. Villa	MFFW	1.632011	-0.309063	0.884961	0.451286	-0.238640	NaN	NaN	-3.302705
16	N. Yamashita	FW	3.322624	-0.187722	-0.846743	-2.499978	-0.193057	NaN	NaN	-2.757956
18	T. Larsson	DF	-0.118982	0.661666	1.418978	1.180790	-0.500742	NaN	NaN	-2.514351
22	B. Katooko	MF	-0.541636	4.180558	-0.372564	-1.259039	-0.432367	NaN	NaN	-2.209709
3	M. Asiimwe	GK	-0.602015	-0.066381	0.185933	1.606172	-1.230068	-1.253662	-0.739575	-7.419922
25	E. Bahrami	DF	-0.843531	0.297643	1.476097	1.606172	-1.218673	NaN	NaN	-2.062101
28	T. Nouri	DF	0.303671	0.054960	1.224955	1.133852	0.034858	NaN	NaN	-2.011924
29	W. Dehghani	MF	-0.964289	0.176301	2.079927	1.608127	-1.207277	NaN	NaN	-1.984217
31	E. Nakanjako	MFFW	0.243292	0.418984	0.964746	0.778879	-0.022121	NaN	NaN	-1.890153
33	K. Al-Zahrani	FW	1.028220	0.540325	-0.260140	-0.580385	-0.329806	NaN	NaN	-1.840628
34	C. Baluka	DF	0.484808	0.297643	0.777976	0.636107	0.183002	NaN	NaN	-1.774382
39	H. Amade	MFFW	0.907462	0.054960	0.340063	0.367188	0.342542	NaN	NaN	-1.663179
42	B. Lai	DF	-0.602015	0.904348	1.069011	0.694780	-0.637490	NaN	NaN	-1.517731
46	H. Zare	DF	-0.783152	0.661666	1.490604	1.333341	-0.272827	NaN	NaN	-1.405101
70	V. Sultan	DFMF	-0.783152	0.661666	0.913974	0.954898	0.103232	NaN	NaN	-0.494604

Surprisingly, Raritan players in the assumed higher leagues seemed to not perform as well as players in the RFL category. Younger players with the lowest values were prioritised due to having more time and potential to train and become strong, homegrown players.

PROBABILITY RANGES OF THE "SUCCESS" BEING COMPETITIVE

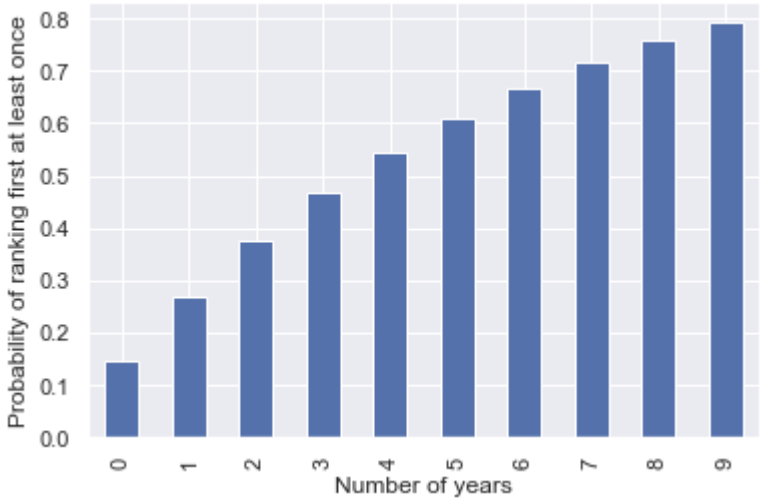
The probabilities of success were calculated by considering the variance of the error terms in the lasso regression (see Appendix C2 for code).

The probability of ranking high should improve over years. As a conservative approach, the probabilities were assumed constant and independent, giving the 5 and 10-year probabilities using a Geometric distribution.

Table 2.3 Probability of reaching ranks

Probability of reaching rank 1 in 1 year	Probability of becoming a top 10 team in 1 year	Probability of becoming a top 10 team in 5 years	Probability of reaching rank 1 in 10 years
14.57%	82.14%	99.98%	79.29%

Figure 2.3 Probability of winning



SPENDING ON ASSEMBLING TEAM

Salaries

The mean salary for Rarita is lower than other nations, likely due to the absence of a national team. The salary projection component was based off the factors:

- Inflation
- Year on year growth rate
- Transition factors increase of league to tournament

$$Salary_t = Salary_{t-1} \cdot (1 + \inf lation_t) + \Pr\left[Rarita's\ rank = 1\right] \cdot Winning\ bonus$$

(See Appendix D for further details).

Table 2.4 Projection of salary

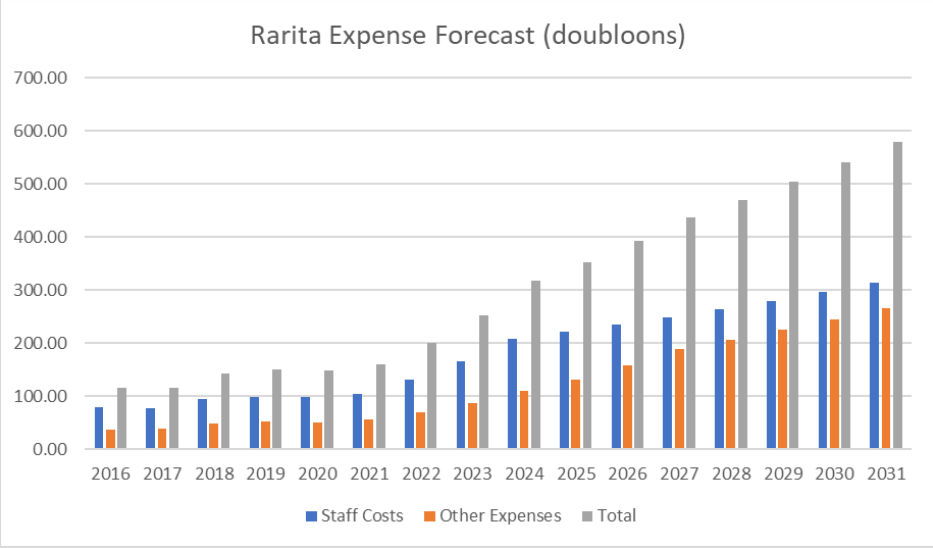
Raritan National Team Salaries (ð millions)												
Player	Pos	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
B. Ampofo	GK	3.84	4.27	4.44	4.61	4.79	4.98	5.17	5.37	5.58	5.80	6.03
B. Katooko	MF	5.66	6.30	6.54	6.80	7.06	7.34	7.62	7.92	8.23	8.55	8.88
B. Lai	DF	15.49	17.23	17.90	18.60	19.32	20.08	20.86	21.67	22.52	23.40	24.31
C. Baluka	DF	7.76	8.63	8.97	9.32	9.68	10.06	10.45	10.86	11.28	11.72	12.18
E. Bahrami	DF	1.25	1.39	1.44	1.50	1.56	1.62	1.68	1.75	1.82	1.89	1.96
E. Nakanjako	MFFW	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.75	11.17	11.60	12.06
M. Asiimwe	GK	4.49	4.99	5.19	5.39	5.60	5.82	6.05	6.28	6.53	6.78	7.05
G. Jankowski	MF	9.02	10.03	10.42	10.83	11.25	11.69	12.15	12.62	13.11	13.63	14.16
H. Amade	MFFW	1.71	1.90	1.98	2.05	2.13	2.22	2.30	2.39	2.49	2.58	2.68
H. Mubaiwa	FW	1.58	1.76	1.83	1.90	1.97	2.05	2.13	2.21	2.30	2.39	2.48
H. Zare	DF	15.43	17.16	17.83	18.53	19.25	20.00	20.78	21.59	22.43	23.31	24.22
J. Namirembe	MF	8.33	9.26	9.63	10.00	10.39	10.80	11.22	11.66	12.11	12.58	13.07
K. Al-Zahrani	FW	6.24	6.94	7.21	7.49	7.78	8.09	8.40	8.73	9.07	9.43	9.79
K. Kazlo?	FW	2.37	2.64	2.74	2.85	2.96	3.07	3.19	0.00	0.00	0.00	0.00
N. Yamashita	FW	8.27	9.20	9.56	9.93	10.32	10.72	11.14	11.57	12.02	12.49	12.98
P. Villa	MFFW	7.82	8.70	9.04	9.39	9.76	10.14	10.53	10.94	11.37	11.81	12.27
T. Larsson	DF	1.14	1.27	1.32	1.37	1.42	1.48	1.54	1.60	1.66	1.72	1.79
T. Nouri	DF	19.54	21.73	22.58	23.46	24.38	25.33	26.31	27.34	28.41	29.52	30.67
V. Sultan	DFMF	5.08	5.65	5.87	6.10	6.34	6.58	6.84	7.11	7.39	7.67	7.97

Expenses

The countries of Galamily, Greri, Nganion and Sobianitedrucy were investigated as benchmarks to project Rarita's expenses, as they have similar GDP to Rarita and high rankings for 2020 and 2021 (Appendix G).

The expense projection assumes a target to emulate the investments of competitive teams. To reach this quota within 5 years, a significant growth rate was apportioned to the first year reflecting the development of the national team, and a uniform increase for the 4 following years. Finally, expenses increase smoothly with stable reinvestments, returning to the average Raritan expense growth rate for 2016-2020.

Figure 2.4 Rarita Expense Forecast



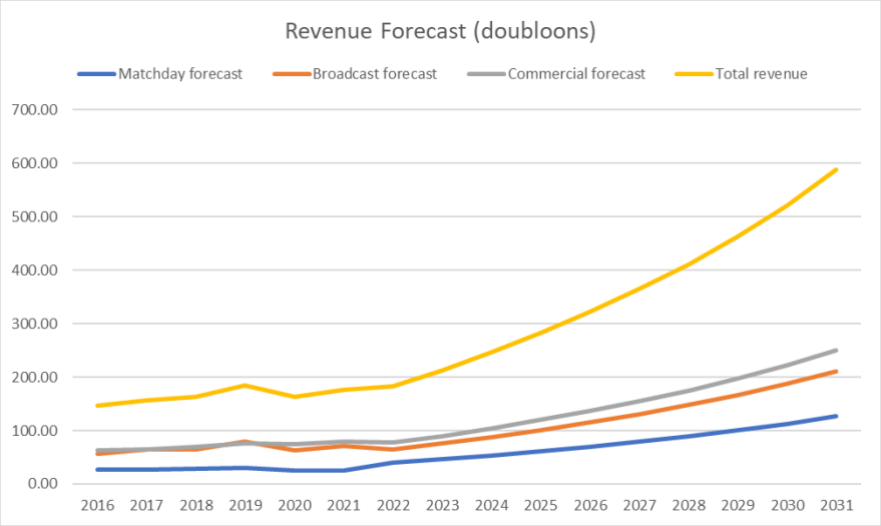
DIRECT TEAM REVENUES

Revenue

Revenue projections exhibit an exponential trend. For the competitively successful countries similar in GDP per capita as Rarita, Commercial revenue is strongly correlated with Social Media Followers, while Matchday revenue is highly correlated with League Attendance and Social Media Followers (Appendix E).

As per the social media followers' and league attendance's projections (Appendix F), the revenue and its subcategories are as follows:

Figure 2.5 Rarita Revenue Forecast



Model assumptions for matchday, broadcast and commercial revenues are in the Assumptions section.

03 ECONOMIC IMPACT

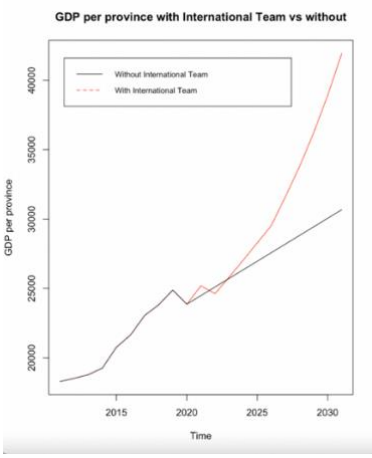
IMPACT ON GDP

The impact of the implementation plan on GDP was modelled through linear regression using Year, Household Savings, Revenue and Expenses as predictors. Healthcare was forecasted using an ARIMA (0,1,0) model with drift, while revenue and expenses were forecasted according to benchmark countries:

$$GDP_t = -445131.51 + 218.49 \times Year_t + 15.75 \times Healthcare_t + 56.69 \times Profit$$

This was used to predict the GDP per capita for 2021 to 2031. Furthermore, GDP was forecasted using a time series with an ARIMA (0,1,0) model with drift to compare the difference in GDP if Rarita does not form a national team. The impact of the implementation plan on GDP is visualised in Figure 3.1, depicting a slight dip in 2022 due to the high growth rate of expenses during this period, yet in 2023 it overtakes the GDP without a national team.

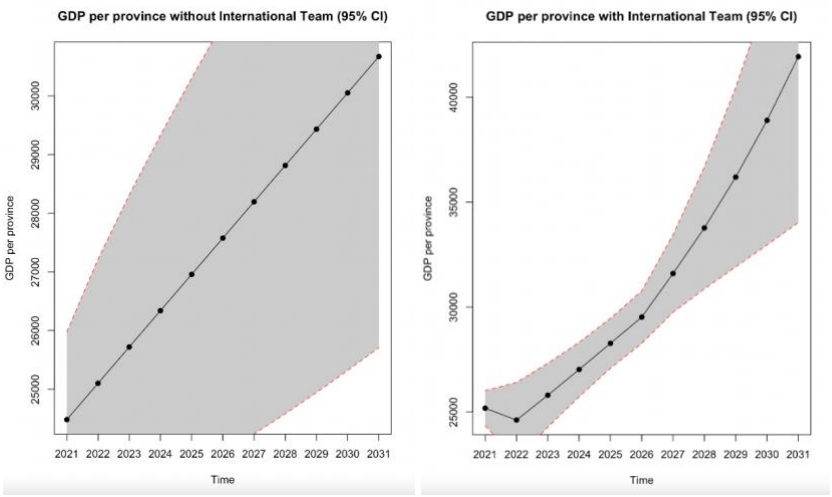
Figure 3.1 GDP per province without tournament and with tournament



By 2031 our implementation plan will lead to a 36.71% increase in GDP per capita in comparison to not implementing an international team (Appendix M and N) due to the positive flow-on effects of a competitive soccer team on the economy.

The introduction of an international soccer team will lead to an influx of new jobs, and according to Okun's law (Appendix O) a countries GDP must grow 4% to achieve a 1% reduction in unemployment. Furthermore, so long as Rarita's heightened growth in GDP exceeds growth in labour productivity, employment and wellbeing index will increase substantially.

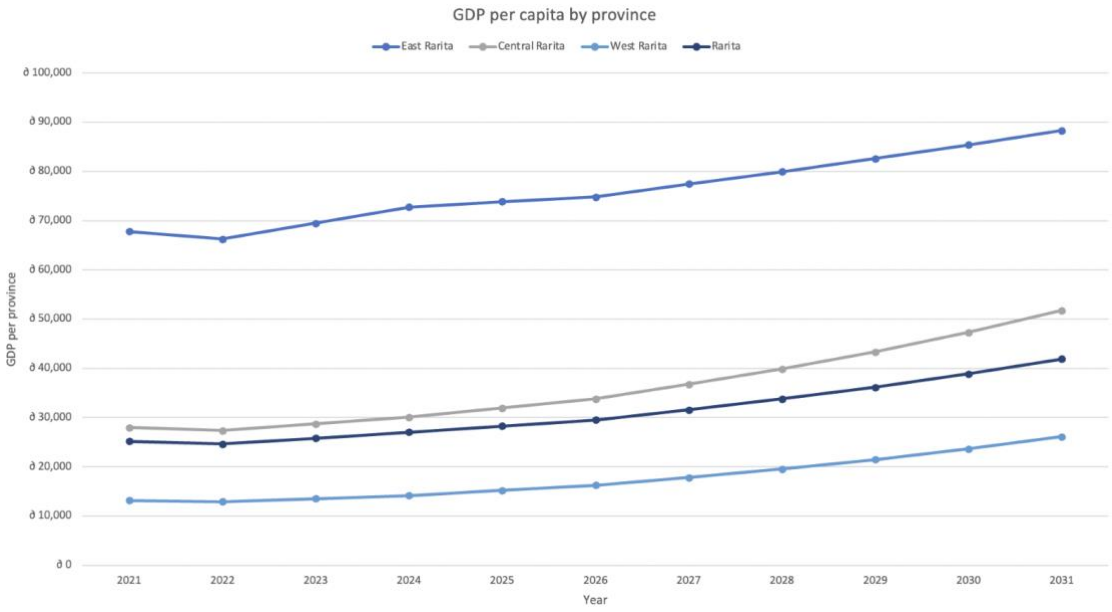
Figure 3.2 GDP per province with tournament and without tournament (95% confidence interval)



IMPACT ON RARITA PROVINCES

East Rarita is the wealthiest province, with a GDP per capita 135% greater than Central Rarita and 410% greater than West Rarita in 2020. Through our implementation plan we aim to lessen the gap by predominately hiring staff from West/Central Rarita, targeting businesses from West/Central Rarita through commercials and as sponsors of our team and by preferencing these locations for future investment projects such as stadiums, soccer schools, merchandise warehouses and TV series. As seen in Figure 3.3, whilst the GDP of all provinces are increasing, Central Rarita and West Rarita have experienced a respective 84.77% and 97.90% increase in GDP per capita from 2021-2031 as opposed to 30.27% for East Rarita (Appendix Q).

Figure 3.3 Impact of GDP per capita on Rarita provinces



IMPACT ON RELATED INDUSTRIES

Related Industries	Impact
Tourism	As a benchmark, the UK is home to the most popular soccer teams in the world, and in 2019 overseas visitors who went to a football match spent 1.4 billion pounds. Ranking in the top 10 will drive inbound tourism to Rarita and increase visitor spending.
Political Influence	Soccer matches have been known to inspire a sense of patriotism in countries.
General Health	The number of individuals playing will rise due to the increased popularity, leading to a heightened level of health/fitness.

04 IMPLEMENTATION PLAN

Net present value of Rarita's chosen team shows a positive impact projected over the next ten years on Rarita's economy and profit. Whilst the first year is expected to be negative, the remaining years are expected to be positive year on year. These values were calculated using the spot rates data and the revenue and expense analysis conducted prior.

Figure 4.1 NPV of Rarita's Chosen Soccer Team

Net Present Value of Rarita's National Soccer Team Project (ð)										
	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
Revenue	192.88	224.15	258.90	297.17	338.96	384.17	432.62	487.18	548.62	617.81
Expense	198.82	222.36	248.70	278.15	311.09	332.83	356.16	381.20	408.06	436.91
Profit	-5.94	1.79	10.20	19.02	27.87	51.34	76.46	105.98	140.56	180.91
Rate	0.14	0.19	0.37	0.61	0.87	1.12	1.33	1.55	1.74	1.90
PV	-5.92	1.78	10.03	18.37	26.27	47.02	67.98	91.00	116.43	144.53
NPV	517.49									

TEAM SELECTION

Twenty players have been selected from the model output, according to age, position, and league, listed in Figure 4.2. The first eighteen constitute the main team and substitutes, and the last two may be assigned once a player retires.

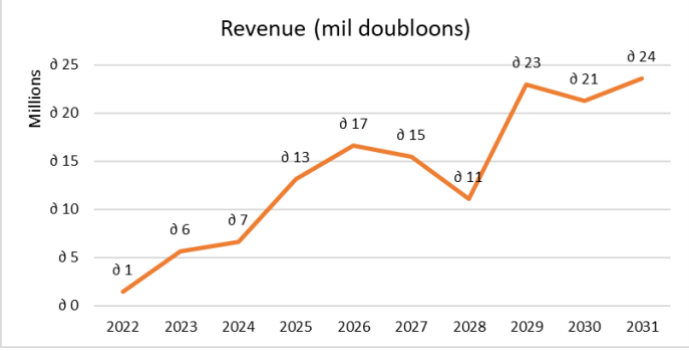
Figure 4.2 Player summary

Player	Age in 2021	Position	League	Model Value
K. Kazlo	38	FW	RFL	-11.3369
B. Ampofo	32	GK	RFL	-8.48254
M. Asiimwe	33	GK	RFL	-7.41992
J. Namirembe	30	MF	RFL	-4.57882
H. Mubaiwa	31	FW	B	-4.26947
G. Jankowski	21	MF	RFL	-3.53385
P. Villa	20	MFFW	RFL	-3.3027
N. Yamashita	19	FW	RFL	-2.75796
T. Larsson	19	DF	RFL	-2.51435
B. Katooko	20	MF	RFL	-2.2071
E. Bahrami	22	DF	RFL	-2.0621
T. Nouri	22	DF	RFL	-2.01192
W. Dehghani	18	MF	RFL	-1.98422
E. Nakanjako	20	MFFW	A	-1.89015
K. Al-Zahrani	18	FW	RFL	-1.84063
C. Baluka	20	DF	RFL	-1.77438
H. Amade	20	MFFW	RFL	-1.66318
B. Lai	23	DF	B	-1.51773
H. Zare	29	DF	D	-1.4051
V. Sultan	20	DFMF	RFL	-0.4946

SOURCES OF REVENUE

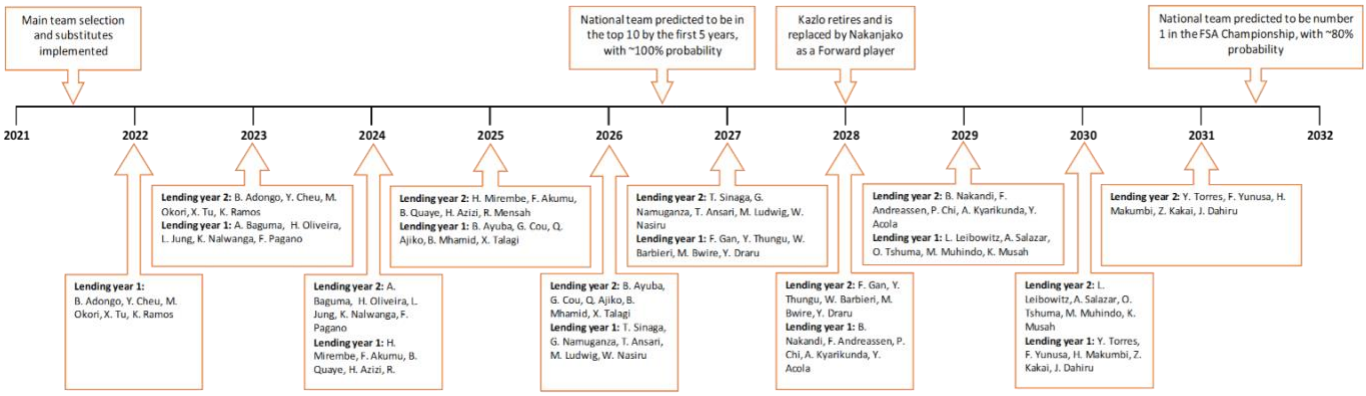
Lending Rarita's under-utilised players is extremely beneficial for the country's revenue as it reduces expenses, keeps players active even though they're not playing for the Raritan team, and develops younger players' athleticism[i]. The total lending revenue figures are:

Figure 4.3 – Appendix K consists of loaned players and revenues



The timeline in Figure 4.4 provides a high-level view of pivotal events for the national team over time, including new player inductions, player lending, and the probabilities of success in 5 and 10 years.

Figure 4.4 Implementation Plan Timeline



REPORTING AND MONITORING PROGRESS

An appropriate reporting frequency to assess progress is every 6 months, and the key metrics we determined are:

- Success rate i.e.: wins/losses, ranking in tournaments
- Year on year growth rate and player statistics for each position
- Participation growth and player pool to monitor the growth of the football “branding” and nation engagement
- Viewership and league attendance
- Entry of new teams which challenge team competitiveness
- Performance statistics of younger players and ensuring they are not “poached” by other national teams
- Social media engagement on platforms such as Facebook and Instagram.
- Potential revenue streams that can increase budget e.g., lending more players, sports brands team sponsorships
- Revenue and expenses should be monitored through an Actual vs Expected analysis

ASSUMPTIONS

GENERAL ASSUMPTIONS

1. The national team consists of 16 players [ii]

Total	Defenders	Midfielders	Forwards	Goalkeeper
11 Main players	4	4	2	1
7 Substitutes	2	2	2	1
2 Replacements for retirees				

2. Substitution of players once they reach the age 45 (max. player age)
3. COVID-19 doesn't exist in this world
4. League attendance projection
   - Assume base 5% increase in annual league attendance
   - Assume 10% increase on the year the national team is created
   - Assume attendance increases dramatically if Rarita wins in any year
5. Social media followers' projection
   - Social media followers for similar GDP benchmark countries have been used to predict Rarita's progression
   - Assume 5% increase in annual social media followers
   - Assume 30% increase in the year the national team is created
   - Assume followers increase dramatically if Rarita wins in any year. Larger increases occur if Rarita wins in earlier years due to the lower likelihood, hence there will be slow decay in the growth factor.

ECONOMIC ASSUMPTIONS

1. Player lending fee remains at 10% of players' salary for the next 10 years.
2. Rarita is allowed to loan players every year
3. Assuming no sanctions are in place for loaning players[iii], 5 players would be a suitable number to loan every year
4. All players loaned can be loaned up to 2 years
5. Revenue modelling is assumed as follows:

$$Matchday_t = \frac{\rho_s}{\rho_s + \rho_a} \cdot revenue\ per\ social\ media \cdot s_t + \frac{\rho_a}{\rho_s + \rho_a} \cdot revenue\ per\ attendance \cdot a_t$$

$$Commercial_t = Commercial_{t-1} \cdot revenue\ per\ social\ media \cdot s_t$$

$$Broadcast_t = Commercial_t \cdot 0.97$$

Where ρ_s, ρ_a are Matchday Revenue's correlation with social media and attendance respectively, and s_t, a_t are social media followers and league attendance at time t.

6. Broadcast revenue was calculated as a fraction of Commercial revenue, as it is correlated to Commercial Revenue.
7. The average broadcast-commercial revenue ratio for years 2016-2020 of 0.97 was chosen as the scaling factor.
8. Cost of capital was chosen to equal the spot rates in 2021, which assumes there are no other projects to consider i.e., the only alternative to building a national team is to invest the money at the spot rates.

06 RISK AND RISK MITIGATION CONSIDERATIONS

Risk Category	Risk factor	Definition	Mitigation
Financial	Loan provision	Lending countries' player fees rise above current fee of 10% of player's salary.	Observe fee rate movements and only borrow players when comparable Rarita players are not available.
Financial	Unexpected changes in Euro to Doubloon conversion rate	Euro currency may be devalued compared to Doubloon, resulting in losses when lending players.	Observe exchange rate movements and only lend players when comparable Rarita players are not available.
Financial	Unexpected rise in inflation rates	Higher inflation rates reduce Doubloon's purchasing power, linking to its ability to invest in facilities and lend players, as well as potentially increasing salaries of the team.	Prioritise team selection and training to ensure prize money is won and investigate inflation hedging strategies if this becomes an increasing risk.
Financial	Direct team revenue may not be enough to cover future expenses	Relying on revenue from match winnings will not be sustainable, especially in the first few years of playing.	Establish other income sources such as player/team sponsorships, and non-governmental funding such as player lending.
Operational	Player injury prior to or during play	Players returning from injury may take a long time to become competitive again, if ever. Players' injury during matches can hinder mobility, strength, and overall team performance.	Have substitute players available and in good health. Provide support in physiotherapy and rehabilitation.
Operational	Political tensions between countries and between Rarita's provinces would deteriorate professional relationships	Tense political environment may affect player exchange, labour laws, taxes, currency value and regulation. In extreme cases, may cause war.	Consult with Rarita's government and local communities. Consider joint ventures or alliances with Raritan companies once team is more established.
Strategic	Building sporting reputation for a completely new team	Football environment is extremely competitive and quite saturated. Over 100 countries involved	Leverage Raritan players that have already played in FSA leagues. Promote players and team through social media and TV series to form online image and fanbase with all generations.
Strategic	Persuading players in higher leagues to join the national team	Highly qualified players may not wish to join a completely new team with no record of accomplishment	Offer competitive salary benefits provided the players are affordable within our budget. Consider slightly lower-ranked players that show potential to improve.

SWOT Analysis for implementation of Rarita's national football team

Strengths	Weaknesses
▪ Existing players with strong personal brand ▪ Effective team leadership and staff ▪ Lower starting salaries for lower-ranked players ▪ Impressive seed capital of 995,000,000 Doubloons	▪ National team formation is happening for the first time ▪ Raritan players' current average league attendance is the lowest ▪ Lack of recognition of some players due to juniority
Opportunities	**Threats**
▪ Community demand for national team ▪ Several non-governmental funding sources ▪ Economic freedom ▪ Available players to lend from other countries ▪ Social media leverage strengthens reputation and fan engagement ▪ Capitalise on increasing tourism rates ▪ Other investment projects funded by prize money, like football stadiums	▪ Civil war within Raritan provinces ▪ Allocated budget is insufficient for compensating multiple higher-ranked players ▪ Increasing regulations from FSA ▪ New and competitive national teams entering the championship ▪ National team's players sign on with other teams that offer them higher salary ▪ Revenue is lower than expected

07 DATA AND DATA LIMITATIONS

- The player data set used for player selection is limited due to many missing values, causing the exclusion of potential candidates.
- The timeline of the data itself is not adequate in measuring past or future performance of the players reliably, or ascertaining whether they are at the beginning, peak, or downturn in their career.
- Likewise, only 2 years of salaries are provided, which is not enough to predict accurate salary increases over the next 10 years.
- Even though the predictors used for team selection translate to current football logic, there is still a trade-off of interpretability vs accuracy.
- The internal and external environments of football are unpredictable, and predictions of economic impact could be completely opposite to the actual outcomes. For instance, new teams may be introduced into the championship, and players may get injured, challenging revenue that can be generated from rank increases.

APPENDIX

A. Raritan player distribution by position

Defender	76
Midfielder	46
Forward	35
Midfielder, Forward	37
Goalkeeper	20
Defender, Midfielder	9

B. Lasso Regression important predictor coefficients and model efficiency

Statistic	Lasso Coefficient
Gls	-1.606882
Standard FK	-0.948223
Total Cmp%	-0.397181
Medium Cmp%	-0.896960
Blocks Pass	0.909493
Performance GA	4.089069
Penalty Kicks PKm	0.933866
R-squared score	**Mean cross validation score**
0.868616	0.778888

This model was the best performing model compared to ordinary lasso regression, as it attained a higher mean CV score and R-squared, meaning more accurate predictions.

C. 1. Python model code for predicting probability of success of being competitive

1. The variance that each predictor contributes to the model was factored to increase accuracy

```
# Calculating variance of each stat
for stat in lasso_coefs.Stat:
    lasso_coefs.loc[lasso_coefs["Stat"] == stat, "Var"] =  np.var(all_normalised[stat+"_x"]
                                                          - all_normalised[stat+"_y"])
lasso_coefs["Var_contributed"] = lasso_coefs["Coef"] ** 2 * lasso_coefs["Var"]
lasso_coefs
```

	Stat	Coef	Var	Var_contributed
0	Gls	-1.606882	1.302647	3.363526
1	Standard FK	-0.948223	1.608775	1.446494
2	Total Cmp%	-0.397181	0.991748	0.156451
3	Medium Cmp%	-0.896960	0.912047	0.733777
4	Blocks Pass	0.909493	1.440967	1.191937
5	Performance GA	4.089069	1.895961	31.701379
6	Penalty Kicks PKm	0.933866	2.156963	1.881099

```
# Standard deviation of prediction
sd = np.sqrt(np.var(y-y_pred) + lasso_coefs["Var_contributed"].sum())
sd
```

```
6.853526852332301
```

2. The success probabilities were calculated here
   Note that the probability of reaching the top 10 teams within the next 5 years is calculated using

$$1 - (1 - \Pr[r \leq 10.5])^5$$

```
# Top 10 in one year
z_rank_10 = (10.5 - rank) / sd
p_top_10 = norm.cdf(z_rank_10)
p_top_10
```

```
0.8214320129240671
```

```
# Top 10 in five years
1 - (1-p_top_10)**5
```

```
0.9998184409025564
```

```
# First in one year
z_rank_1 = (1.5 - rank) / np.std(y-y_pred)
p_first = norm.cdf(z_rank_1)
p_first
```

```
0.14570505745874235
```

```
# First in ten years
1 - (1-p_first)**10
```

```
0.7929485132643551
```

C. 2. Lasso regression model assumptions

The Lasso Cross Validation model assumes splitting the dataset into 3 groups, as it generates a less biased estimate of performance[iv]. 2 groups are used as training samples, and 1 as a testing sample.

The dependent variable is the team rank, and the independent variables are the player statistics. Error terms are assumed to be normally distributed with mean 0 and constant variance, independent of predictors.

Let r denote the rank and $r*$ denote the predicted rank. Correcting for continuous vs discrete, the probability of ranking first is:

$$\Pr[r \leq 1.5] = \Pr\left[\frac{r - r*}{\sigma_r} \leq \frac{1.5 - r*}{\sigma_r}\right] = \Phi\left(\frac{1.5 - r*}{\sigma_r}\right)$$

Where Φ is the CDF of a standard normal random variable.

The variance σ^2 is estimated from the data on $r - r*$.

The probability of reaching the top 10 is similarly calculated, using $\Pr[r \leq 10.5]$.

D. 1. Characteristics of Salary by Country

Salary by Country

	2020 GDP	Average GDP growth	Social Media	2020 Ranking	2020 mean salary	2021 Ranking	2021 mean salary	Change in mean salary	Average change in salary
Bernepamar	3731	1.145	37.4	14	22967500	8	16940000	0.738	0.795
Byasier Pujan	52450	1.034	75.9	6	19223571.43	15	20091029.41	1.045	1.039
Cuandbo	59329	0.995	8.8	16	17754000	#N/A	22375000	1.260	1.353
Djipines	30498	0.997	2	#N/A	18457857.14	16	17296923.08	0.937	1.079
Dosqaly	39069	1.014	89.6	1	18949886.04	9	20535220.78	1.084	1.190
Eastern Sleboube	10137	1.046	3.1	#N/A	28796666.67	19	24482000	0.850	1.071
Esia	46255	1.024	87.1	9	20682403.85	14	21064310.34	1.018	1.166
Galamily	22955	1.056	7.9	#N/A	18440588.24	7	21512307.69	1.167	0.999
Giumle Lizeibon	48635	1.019	9.4	#N/A	20759411.76	10	22008787.88	1.060	1.176
Greri Landmoslands	22198	1.028	91.7	5	17666600	11	18401607.14	1.042	0.963
Manlisgamncent	14148	1.033	33	12	21048461.54	13	22648620.69	1.076	1.190
Mico	61124	1.029	93.2	7	19054193.55	4	21076296.3	1.106	0.984
Nganion	27090	1.007	248	2	18419974.55	3	19417402.91	1.054	1.094
Nkasland Cronestan	15737	1.061	37.6	10	17781923.08	22	17438571.43	0.981	1.082
People's Land of Maneau	41965	0.998	140.8	8	18469888.89	2	19498965.52	1.056	1.095
Quewenia	87184	1.012	76.4	11	20159666.67	5	21415757.58	1.062	1.532
Rarita	23863	1.025	35.9	#N/A	8052641.509	#N/A	7653859.649	0.950	0.761
Sobianitedrucy	31746	1.008	102.9	3	18875156.25	1	20930227.27	1.109	1.209
Southern Ristan	45205	1.019	251.5	4	20065384.62	6	21415609.76	1.067	0.961
Unicorporated Tiagascar	6086	1.052	4.7	15	18743750	#N/A	19468750	1.039	0.997
Xikong	52327	1.002	31.6	13	21051250	12	22198823.53	1.055	1.028

Here, Raritan salary is distinctly lower than the other countries. This is implied to be due to the lack of a national team.

D. 2. Change in Salary by entering a national team

Impact of transitioning from league to tournament on Salary			
	2020 mean salary	2021 mean salary	Change
Players in 2021, but not 2020 tournament	20037188	21696765	1.083
Players in both 2020 and 2021 tournament	21534111	21782752	1.012
			1.070

The table shows the derivation of the salary increase on the year Rarita creates a national team. The average salary of all players fitting the condition was taken as a proxy.

D. 3. Change in Salary by winning

Comparison of Salary for winning team		
2021>rank 1	2021<rank 1	Change
21520185.87	23858400	1.108652134

The table shows the derivation of salary increase if Rarita wins the tournament. This involved comparing the average salary in 2021 between the winning nation and the rest.

D. 4. Salary Projections

Player	Pos	2020 Salary	2021 Salary	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
B. Ampofo	GK	#N/A	3840000	4270919	4437506	4610592	4790428	4977279	5171418	5373130	5582710	5800464	6026711
B. Katooko	MF	#N/A	5660000	6295156	6540699	6795820	7060892	7336302	7622455	7919770	8228681	8549642	8883121
B. Lai	DF	14760000	15490000	17228263	17900253	18598454	19323889	20077619	20860748	21674423	22519836	23398224	24310874
C. Baluka	DF	#N/A	7760000	8630815	8967461	9317237	9680657	10058252	10450575	10858200	11281725	11721770	12178979
E. Bahrami	DF	#N/A	1250000	1390273	1444501	1500844	1559384	1620208	1683404	1749066	1817288	1888172	1961820
E. Nakanjako	MFFW	9790000	0	0	0	0	0	0	0	10750000	11169305	11604964	12057617
M. Asiimwe	GK	#N/A	4490000	4993861	5188647	5391030	5601308	5819787	6046789	6282644	6527699	6782313	7046858
G. Jankowski	MF	#N/A	9020000	10032210	10423517	10830088	11252516	11691422	12147447	12621259	13113552	13625047	14156493
H. Amade	MFFW	6650000	1710000	1901894	1976077	2053154	2133238	2216445	2302897	2392722	2486050	2583019	2683770
H. Mubaiwa	FW	34860000	1580000	1757305	1825849	1897066	1971062	2047943	2127823	2210819	2297052	2386649	2479741
H. Zare	DF	15280000	15430000	17161530	17830917	18526414	19249038	19999849	20779945	21590468	22432606	23307592	24216707
J. Namirembe	MF	#N/A	8330000	9264780	9626153	10001622	10391736	10797067	11218207	11655774	12110409	12582777	13073569
K. Al-Zahrani	FW	#N/A	6240000	6940243	7210948	7492211	7784446	8088079	8403555	8731336	9071903	9425753	9793406
K. Kazlo?	FW	9470000	2370000	2635958	2738773	2845600	2956592	3071915	3191735	0	0	0	0
N. Yamashita	FW	23040000	8270000	9198046	9556817	9929581	10316886	10719297	11137404	11571819	12023179	12492144	12979401
P. Villa	MFFW	35290000	7820000	8697548	9036797	9389278	9755507	10136022	10531378	10942156	11368955	11812402	12273146
T. Larsson	DF	4160000	1140000	1267929	1317385	1368769	1422158	1477630	1535265	1595148	1657367	1722013	1789180
T. Nouri	DF	19220000	19540000	21732748	22580436	23461188	24376293	25327093	26314978	27341396	28407850	29515901	30667171
V. Sultan	DFMF	#N/A	5080000	5650070	5870451	6099429	6337337	6584526	6841356	7108203	7385459	7673530	7972837
W. Dehghani	MF	#N/A	360000	400399	416016	432243	449103	466620	484820	503731	523379	543793	565004
		Sum	125380000	139449946.5	144889203.9	150540620.1	156412470.3	162513352.5	168852200.1	182872066.4	190005007.5	197416169.5	205116404.6
		% of initial	0.12601005	0.1401507	0.14561729	0.151297106	0.157198463	0.163330003	0.169700704	0.183791021	0.190959806	0.198408211	0.20614714
		Cumulative	0.12601005	0.26616075	0.411778041	0.563075146	0.720273609	0.883603611	1.053304315	1.237095337	1.428055143	1.626463354	1.832610494
		Prob of winning	0	0.145705057	0.145705057	0.145705057	0.145705057	0.145705057	0.145705057	0.145705057	0.145705057	0.145705057	0.145705057
		Winning bonus	1	1.108652134	1.108652134	1.108652134	1.108652134	1.108652134	1.108652134	1.108652134	1.108652134	1.108652134	1.108652134
		Inflation	1.0228	1.0228	1.0228	1.0228	1.0228	1.0228	1.0228	1.0228	1.0228	1.0228	1.0228
		Increase due to new national team	1	1.070464852	1	1	1	1	1	1	1	1	1

- Salary was projected using each component as shown in the table.
- The 'Sum' row reflects the total salaries payable every year.
- As assumed, once Kazlo retires, his salary from 2028 onwards goes to 0, and Nakanjako's salary starts being paid, starting growth from the 2020 salary number.

D. 5. Salary projections of loaned players as a source of revenue

To interpret this, the red numbers are the revenues received each year for each player. For example, B. Adongo is lent for 2022 and 2023, F. Gan is lent for 2027 and 2028. Revenue figures at the bottom are the sum of incoming cashflows for each year.

	Lending price projection (Salary * 10%)									
Player lended	**2022**	**2023**	**2024**	**2025**	**2026**	**2027**	**2028**	**2029**	**2030**	**2031**
B. Adongo	**ð 274,446**	**ð 304,599**	ð 338,065	ð 375,209	ð 416,433	ð 462,186	ð 512,967	ð 569,326	ð 631,878	ð 701,303
Y. Cheu	**ð 735,420**	**ð 816,220**	ð 905,898	ð 1,005,429	ð 1,115,896	ð 1,238,499	ð 1,374,573	ð 1,525,597	ð 1,693,215	ð 1,879,248
A. Baguma	ð 630,869	**ð 700,182**	**ð 777,111**	ð 862,493	ð 957,255	ð 1,062,428	ð 1,179,157	ð 1,308,711	ð 1,452,499	ð 1,612,085
H. Oliveira	ð 323,157	**ð 358,662**	**ð 398,068**	ð 441,804	ð 490,345	ð 544,219	ð 604,013	ð 670,376	ð 744,030	ð 825,776
M. Okori	**ð 1,146,494**	**ð 1,272,459**	ð 1,412,264	ð 1,567,430	ð 1,739,643	ð 1,930,778	ð 2,142,913	ð 2,378,354	ð 2,639,664	ð 2,929,684
H. Mirembe	ð 835,218	ð 926,983	**ð 1,028,831**	**ð 1,141,869**	ð 1,267,326	ð 1,406,567	ð 1,561,106	ð 1,732,625	ð 1,922,988	ð 2,134,267
L. Jung	ð 3,061,674	**ð 3,398,060**	**ð 3,771,405**	ð 4,185,769	ð 4,645,659	ð 5,156,078	ð 5,722,576	ð 6,351,315	ð 7,049,134	ð 7,823,622
B. Ayuba	ð 2,607,829	ð 2,894,351	ð 3,212,353	**ð 3,565,294**	**ð 3,957,013**	ð 4,391,770	ð 4,874,293	ð 5,409,832	ð 6,004,210	ð 6,663,892
X. Tu	**ð 2,115,965**	**ð 2,348,446**	ð 2,606,469	ð 2,892,842	ð 3,210,679	ð 3,563,436	ð 3,954,951	ð 4,389,481	ð 4,871,753	ð 5,407,013
F. Akumu	ð 665,323	ð 738,422	**ð 819,552**	**ð 909,597**	ð 1,009,534	ð 1,120,452	ð 1,243,555	ð 1,380,185	ð 1,531,826	ð 1,700,127
F. Gan	ð 520,378	ð 577,552	ð 641,007	ð 711,435	ð 789,600	**ð 876,353**	**ð 972,638**	ð 1,079,502	ð 1,198,107	ð 1,329,743
K. Ramos	**ð 116,432**	**ð 129,224**	ð 143,422	ð 159,179	ð 176,668	ð 196,079	ð 217,622	ð 241,532	ð 268,070	ð 297,522
G. Cou	ð 944,521	ð 1,048,296	ð 1,163,472	**ð 1,291,302**	**ð 1,433,178**	ð 1,590,641	ð 1,765,405	ð 1,959,370	ð 2,174,646	ð 2,413,574
K. Nalwanga	ð 607,107	**ð 673,810**	**ð 747,842**	ð 830,007	ð 921,200	ð 1,022,412	ð 1,134,744	ð 1,259,419	ð 1,397,791	ð 1,551,366
F. Pagano	ð 484,735	**ð 537,993**	**ð 597,102**	ð 662,706	ð 735,518	ð 816,329	ð 906,019	ð 1,005,563	ð 1,116,045	ð 1,238,664
Y. Thungu	ð 3,097,316	ð 3,437,618	ð 3,815,309	ð 4,234,497	ð 4,699,741	**ð 5,216,102**	**ð 5,789,195**	ð 6,425,254	ð 7,131,196	ð 7,914,701
B. Quaye	ð 2,576,939	ð 2,860,067	**ð 3,174,302**	**ð 3,523,063**	ð 3,910,142	ð 4,339,749	ð 4,816,557	ð 5,345,752	ð 5,933,090	ð 6,584,958
L. Leibowitz	ð 1,021,746	ð 1,134,005	ð 1,258,598	ð 1,396,881	ð 1,550,356	ð 1,720,693	ð 1,909,746	**ð 2,119,570**	**ð 2,352,447**	ð 2,610,910
H. Azizi	ð 697,401	ð 774,025	**ð 859,067**	**ð 953,452**	ð 1,058,208	ð 1,174,473	ð 1,303,513	ð 1,446,730	ð 1,605,682	ð 1,782,098
R. Mensah	ð 629,681	ð 698,864	**ð 775,648**	**ð 860,868**	ð 955,452	ð 1,060,427	ð 1,176,936	ð 1,306,246	ð 1,449,764	ð 1,609,049
Q. Ajiko	ð 4,628,747	ð 5,137,308	ð 5,701,743	**ð 6,328,194**	**ð 7,023,472**	ð 7,795,141	ð 8,651,593	ð 9,602,144	ð 10,657,131	ð 11,828,030
A. Salazar	ð 2,718,320	ð 3,016,981	ð 3,348,457	ð 3,716,352	ð 4,124,668	ð 4,577,845	ð 5,080,812	**ð 5,639,041**	**ð 6,258,603**	ð 6,946,235
B. Mhamid	ð 873,236	ð 969,179	ð 1,075,663	**ð 1,193,846**	**ð 1,325,013**	ð 1,470,593	ð 1,632,167	ð 1,811,493	ð 2,010,521	ð 2,231,417
X. Takagi	ð 571,465	ð 634,252	ð 703,937	**ð 781,279**	**ð 867,118**	ð 962,388	ð 1,068,125	ð 1,185,480	ð 1,315,729	ð 1,460,288
Y. Torres	ð 3,124,642	ð 3,467,946	ð 3,848,970	ð 4,271,856	ð 4,741,204	ð 5,262,120	ð 5,840,270	ð 6,481,940	**ð 7,194,110**	**ð 7,984,527**
T. Sinaga	ð 2,750,398	ð 3,052,584	ð 3,387,971	ð 3,760,208	**ð 4,173,342**	**ð 4,631,867**	ð 5,140,770	ð 5,705,586	ð 6,332,458	ð 7,028,206
O. Tshuma	ð 3,332,555	ð 3,698,703	ð 4,105,080	ð 4,556,105	ð 5,056,684	ð 5,612,262	ð 6,228,881	**ð 6,913,248**	**ð 7,672,806**	ð 8,515,817
M. Muhindo	ð 3,402,652	ð 3,776,501	ð 4,191,425	ð 4,651,937	ð 5,163,045	ð 5,730,309	ð 6,359,898	**ð 7,058,660**	**ð 7,834,195**	ð 8,694,938
F. Yunusa	ð 2,156,359	ð 2,393,279	ð 2,656,228	ð 2,948,068	ð 3,271,972	ð 3,631,463	ð 4,030,452	ð 4,473,278	**ð 4,964,757**	**ð 5,510,235**
H. Makumbi	ð 882,741	ð 979,728	ð 1,087,370	ð 1,206,840	ð 1,339,435	ð 1,486,599	ð 1,649,932	ð 1,831,210	**ð 2,032,405**	**ð 2,255,705**
Z. Kakai	ð 820,961	ð 911,160	ð 1,011,269	ð 1,122,377	ð 1,245,693	ð 1,382,557	ð 1,534,459	ð 1,703,050	**ð 1,890,164**	**ð 2,097,836**
J. Dahiru	ð 2,268,039	ð 2,517,228	ð 2,793,796	ð 3,100,750	ð 3,441,429	ð 3,819,539	ð 4,239,192	ð 4,704,952	**ð 5,221,885**	**ð 5,795,613**
B. Nakandi	ð 1,135,801	ð 1,260,592	ð 1,399,093	ð 1,552,811	ð 1,723,419	ð 1,912,771	**ð 2,122,927**	**ð 2,356,173**	ð 2,615,045	ð 2,902,360
F. Andreassen	ð 666,511	ð 739,741	ð 821,016	ð 911,221	ð 1,011,337	ð 1,122,452	**ð 1,245,776**	**ð 1,382,650**	ð 1,534,561	ð 1,703,163
P. Chi	ð 1,498,165	ð 1,662,768	ð 1,845,456	ð 2,048,217	ð 2,273,254	ð 2,523,017	**ð 2,800,221**	**ð 3,107,881**	ð 3,449,343	ð 3,828,323
A. Kyarikunda	ð 1,053,824	ð 1,169,608	ð 1,298,113	ð 1,440,736	ð 1,599,030	ð 1,774,715	**ð 1,969,703**	**ð 2,186,114**	ð 2,426,303	ð 2,692,880
K. Musah	ð 615,424	ð 683,040	ð 758,086	ð 841,377	ð 933,819	ð 1,036,418	ð 1,150,289	**ð 1,276,671**	**ð 1,416,939**	ð 1,572,618
W. Barbieri	ð 192,468	ð 213,615	ð 237,085	ð 263,133	ð 292,044	**ð 324,131**	**ð 359,743**	ð 399,268	ð 443,135	ð 491,823
M. Bwire	ð 4,012,135	ð 4,452,949	ð 4,942,194	ð 5,485,193	ð 6,087,851	**ð 6,756,723**	**ð 7,499,084**	ð 8,323,008	ð 9,237,457	ð 10,252,376
G. Namuganza	ð 3,523,836	ð 3,911,000	ð 4,340,701	ð 4,817,614	**ð 5,346,925**	**ð 5,934,391**	ð 6,586,403	ð 7,310,051	ð 8,113,206	ð 9,004,604
T. Ansari	ð 635,621	ð 705,457	ð 782,965	ð 868,990	**ð 964,466**	**ð 1,070,431**	ð 1,188,040	ð 1,318,570	ð 1,463,441	ð 1,624,229
M. Ludwig	ð 1,414,999	ð 1,570,465	ð 1,743,012	ð 1,934,517	**ð 2,147,063**	**ð 2,382,960**	ð 2,644,776	ð 2,935,358	ð 3,257,865	ð 3,615,807
W. Nasiru	ð 2,662,480	ð 2,955,007	ð 3,279,673	ð 3,640,011	**ð 4,039,939**	**ð 4,483,807**	ð 4,976,443	ð 5,523,204	ð 6,130,039	ð 6,803,546
Y. Acola	ð 1,587,271	ð 1,761,664	ð 1,955,218	ð 2,170,038	ð 2,408,460	ð 2,673,077	**ð 2,966,768**	**ð 3,292,727**	ð 3,654,499	ð 4,056,018
Y. Draru	ð 1,353,219	ð 1,501,898	ð 1,666,911	ð 1,850,055	ð 2,053,320	**ð 2,278,918**	**ð 2,529,303**	ð 2,807,198	ð 3,115,624	ð 3,457,938
Revenue (doubloons)	**ð 4,388,756**	**ð 5,668,708**	**ð 6,657,400**	**ð 13,159,915**	**ð 16,671,733**	**ð 15,452,227**	**ð 11,105,395**	**ð 23,007,190**	**ð 21,303,320**	**ð 23,643,916**

E. Correlation matrix for revenue subcategories

	Matchday	Broadcast	Commercial
Attendance	0.322769	-0.193377	0.295109
Social media	0.537578	0.154145	0.818948

- Moderately high correlation between Matchday Revenue and social media.
- Commercial Revenue strongly correlated with social media.
- Broadcast shows weak correlation with both factors.

F. Social media followers and league attendance projections

Rarita Social Media Forecast										
2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
35.9	52.10986	60.69437	70.22882	80.72383	92.16974	104.5335	117.7563	132.6515	149.431	168.3329
Growth rates										
Prob win	0.1457	0.1457	0.1457	0.1457	0.1457	0.1457	0.1457	0.1457	0.1457	0.1457
Win	1.8	1.75	1.7	1.65	1.6	1.55	1.5	1.5	1.5	1.5
Base	1.3	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05

Matchday Revenue per attendance/ social media		
	Attendance	**Social Media**
Galamily	0.0009	5.4937
Greri Landmoslands	0.0009	0.5387
Nganion	0.0011	0.3147
Sobianitedrucy	0.0012	0.4499
Average (Ignoring outliers)	0.0010	0.4344

Derivation of scaling factor for Matchday Revenue using Attendance and Social Media.

Rarita League Attendance										
2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
28,067.00	32,897.93	36,807.63	41,181.97	46,076.17	51,552.01	57,678.62	64,533.33	72,202.69	80,783.49	90,384.06
Growth Rates										
Prob win	0.1457	0.1457	0.1457	0.1457	0.1457	0.1457	0.1457	0.1457	0.1457	0.1457
Win	1.45	1.45	1.45	1.45	1.45	1.45	1.45	1.45	1.45	1.45
Base	1.1	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05

Commercial Revenue per attendance/ social media		
	Attendance	**Social Media**
Galamily	0.0013	7.6911
Greri Landmoslands	0.0027	1.5772
Nganion	0.0029	0.8470
Sobianitedrucy	0.0052	2.0246
Average (Ignoring outliers)	0.0030	1.4829

Derivation of scaling factor for Commercial Revenue using Social Media.

G. Comparison of Rarita's expenses with other countries

Country	2016	2017	2018	2019	2020	2020 GDP differential with Rarita
Bernepamar	ð 2,190	ð 2,643	ð 3,100	ð 3,666	ð 3,731	20132
Byasier Pujan	ð 46,054	ð 48,604	ð 53,072	ð 52,529	ð 52,450	28587
Cuandbo	ð 62,050	ð 72,082	ð 74,544	ð 69,010	ð 59,329	35466
Djipines	ð 31,034	ð 30,333	ð 32,602	ð 31,975	ð 30,498	6635
Dosqaly	ð 37,074	ð 38,724	ð 41,614	ð 40,619	ð 39,069	15206
Eastern Sleboube	ð 8,714	ð 10,731	ð 11,299	ð 11,509	ð 10,137	13726
Esia	ð 42,150	ð 44,587	ð 47,998	ð 46,842	ð 46,255	22392
Galamily	ð 18,594	ð 20,657	ð 23,443	ð 23,684	ð 22,955	908
Giumle Lizeibon	ð 45,322	ð 47,359	ð 51,513	ð 50,165	ð 48,635	24772
Greri Landmoslands	ð 19,998	ð 21,459	ð 23,575	ð 23,354	ð 22,198	1665
Manlisgamncent	ð 12,540	ð 13,643	ð 15,243	ð 15,327	ð 14,148	9715
Mico	ð 54,719	ð 57,668	ð 61,654	ð 59,836	ð 61,124	37261
Nganion	ð 26,532	ð 28,129	ð 30,380	ð 29,585	ð 27,090	3227
Nkasland Cronestan	ð 12,460	ð 13,879	ð 15,484	ð 15,748	ð 15,737	8126
People's Land of Maneau	ð 42,473	ð 41,720	ð 44,504	ð 43,969	ð 41,965	18102
Quewenia	ð 83,156	ð 83,435	ð 86,475	ð 85,420	ð 87,184	63321
Sobianitedrucy	ð 30,971	ð 32,359	ð 34,640	ð 33,675	ð 31,746	7883
Southern Ristan	ð 42,026	ð 44,133	ð 47,567	ð 46,638	ð 45,205	21342
Unicorporated Tiagascar	ð 5,000	ð 5,400	ð 6,076	ð 6,126	ð 6,086	17777
Xikong	ð 52,017	ð 53,845	ð 54,644	ð 51,991	ð 52,327	28464

Country	2020 rank	2021 rank
Galamily	Unranked	7
Greri Landmoslands	5	11
Nganion	2	3
Sobianitedrucy	3	1

Gamily, Greri Landmoslands, Nganion and Sobianitedrucy had similar GDP per capita as Rarita and are competitively successful, thus used as a benchmark for expenses and revenue projections.

H. Expense breakdown of benchmark countries based on 2016-2020 data (doubloons)

Figures highlighted in yellow are the expense targets to be reached in the first 3 years

Total Expenses (1) + (2)					
	2016	**2017**	**2018**	**2019**	**2020**
Galamily	213.41	231.32	295.03	284.23	276.62
Greri Landmoslands	196.38	178.13	196.63	258.05	272.54
Nganion	315.36	295.58	453.06	449.75	435.28
Sobianitedrucy	259.58	313.62	342.01	418.63	361.98
Average	*246.18*	*254.66*	*321.68*	*352.67*	*336.61*
Staff costs (1)					
Galamily	137.75	133.09	153.94	150.42	164.4
Greri Landmoslands	155.94	125.46	144.58	188.68	188.79
Nganion	252.69	232.49	317.6	300.64	303.42
Sobianitedrucy	195.63	227.3	246.2	301.54	246.11
Average	*185.5*	*179.59*	*215.58*	*235.32*	***225.68***
Other costs (2)					
Galamily	75.66	98.23	141.09	133.81	112.22
Greri Landmoslands	40.44	52.67	52.05	69.37	83.75
Nganion	62.67	63.09	135.46	149.11	131.86
Sobianitedrucy	63.95	86.32	95.81	117.09	115.87
Average	*60.68*	*75.08*	*106.1*	*117.35*	***110.93***

I. Average staff costs and other costs growth rates based on 2016-2020 data

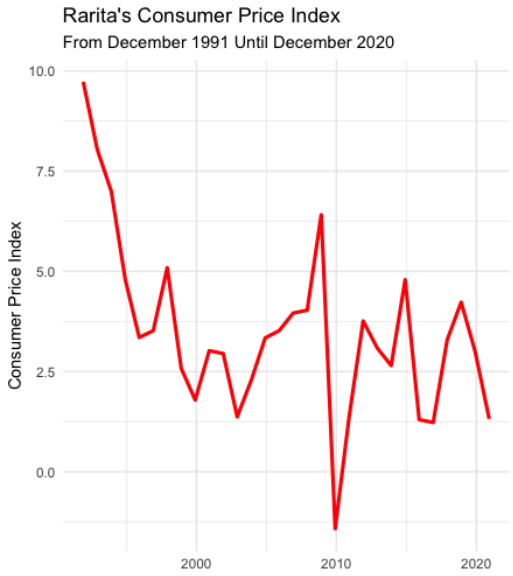
Country	Staff costs growth rate
Galamily	1.048226
Greri	1.065636
Nganion	1.060497
Sobianitedrucy	1.071498
Average	*1.061464*

Country	Other costs growth rate
Galamily	1.130421
Greri	1.207676
Nganion	1.284719
Sobianitedrucy	1.167858
Average	*1.197668*

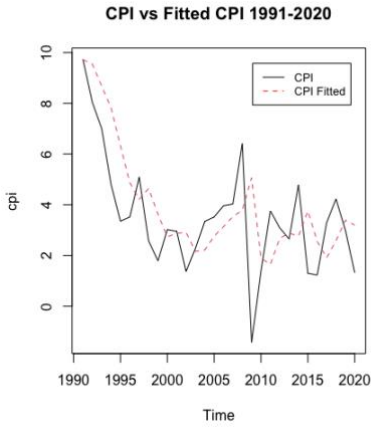
J. Economic projections for Rarita

Inflation and household savings were initially considered as predictors, however in order to do so they were forcasted from 2021-2031 using a time series with an ARIMA(0,1,1) and ARIMA(0,1,0) model respectively which output constant values for all years. Thus, they were not considered. Results from this can be seen below;
Inflation (Time Series)

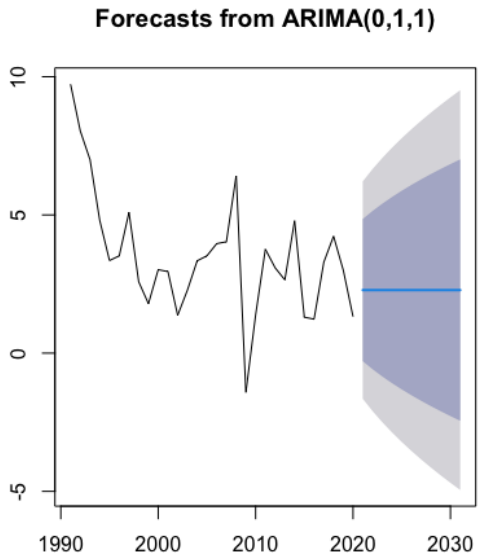
J. 1. Consumer Price Index



J. 2. CPI vs Fitted CPI for 199-2020 using ARIMA(0,1,1) model



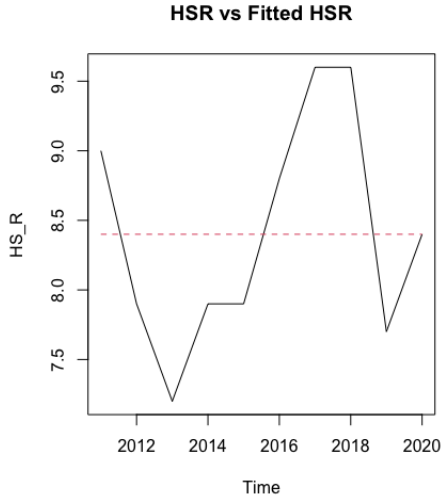
J. 3. Forecasted CPI for 2021-2031 using an ARIMA(0,1,1) model



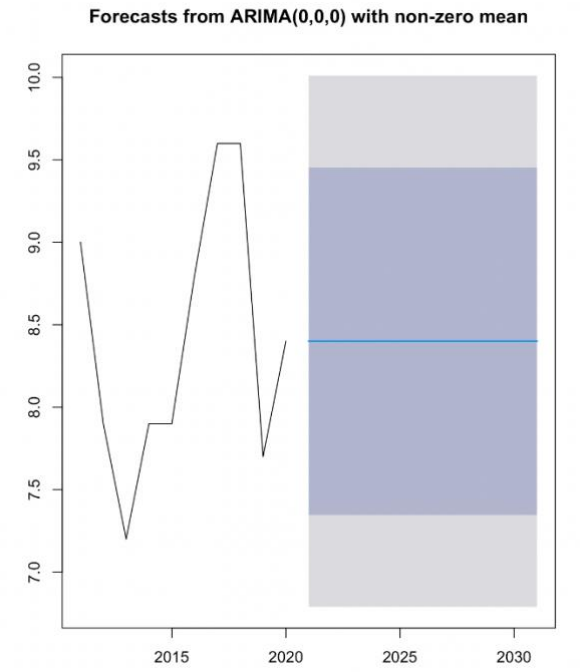
J. 4. R output showing CPI forecast for 2021-2031 including 90% and 95% confidence interval

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	2.281277	-0.2890311	4.851586	-1.649670	6.212225
2022	2.281277	-0.5789435	5.141498	-2.093053	6.655607
2023	2.281277	-0.8420609	5.404615	-2.495456	7.058011
2024	2.281277	-1.0846727	5.647227	-2.866499	7.429053
2025	2.281277	-1.3109362	5.873491	-3.212539	7.775094
2026	2.281277	-1.5237688	6.086323	-3.538038	8.100593
2027	2.281277	-1.7253115	6.287866	-3.846271	8.408826
2028	2.281277	-1.9171905	6.479745	-4.139725	8.702280
2029	2.281277	-2.1006755	6.663230	-4.420341	8.982896
2030	2.281277	-2.2767802	6.839335	-4.689670	9.252225
2031	2.281277	-2.4463295	7.008884	-4.948973	9.511528

K. Household Savings (Time Series)



Household Savings Rate vs Fitted Household Savings Rate for 2011-2020 using an ARIMA(0,0,0) model with a non-zero mean



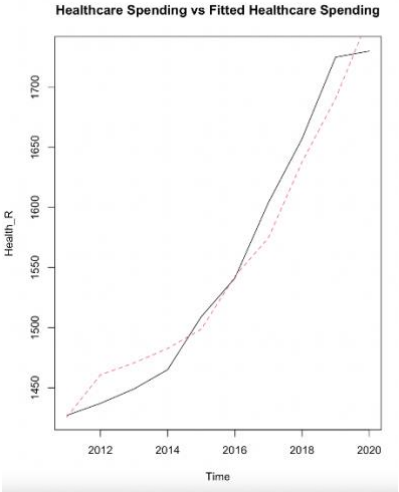
Forecasted Household Savings Rate for 2021-2031 using an ARIMA(0,0,0) model with a non-zero mean

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	8.4	7.346665	9.453335	6.789063	10.01094
2022	8.4	7.346665	9.453335	6.789063	10.01094
2023	8.4	7.346665	9.453335	6.789063	10.01094
2024	8.4	7.346665	9.453335	6.789063	10.01094
2025	8.4	7.346665	9.453335	6.789063	10.01094
2026	8.4	7.346665	9.453335	6.789063	10.01094
2027	8.4	7.346665	9.453335	6.789063	10.01094
2028	8.4	7.346665	9.453335	6.789063	10.01094
2029	8.4	7.346665	9.453335	6.789063	10.01094
2030	8.4	7.346665	9.453335	6.789063	10.01094
2031	8.4	7.346665	9.453335	6.789063	10.01094

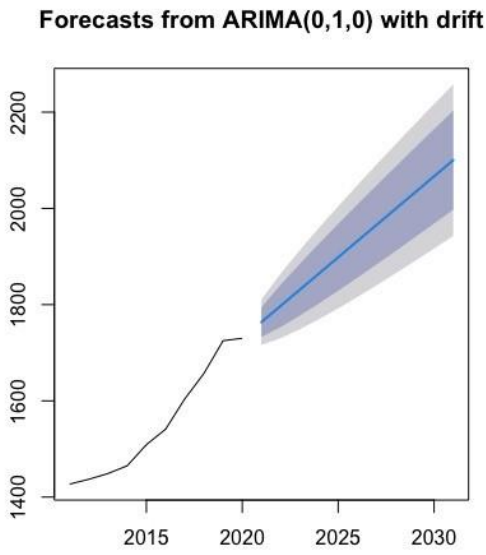
R output showing household savings rate forecast for 2021-2031 with 80% and 95% confidence intervals

L. Healthcare Spending (Time Series)

Healthcare spending was a significant predictor used in the linear regression model predicting GDP per province from 2021-2031. It was forecasted using as a time series using an ARIMA (0,1,0) model with drift.



Healthcare Spending vs Fitted Healthcare Spending for 2011-2020 using an ARIMA(0,1,0) model with drift

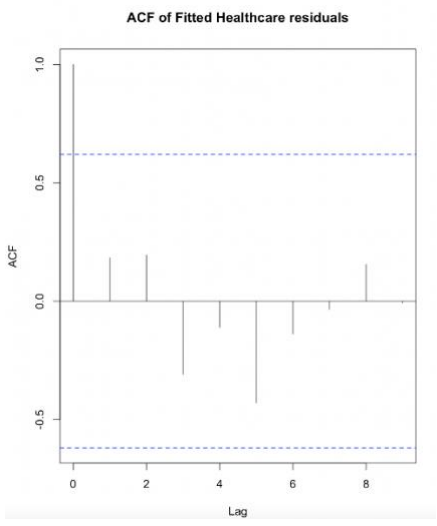


Forecasted Healthcare Spending for 2021-2031 using an ARIMA(0,1,0) model with drift

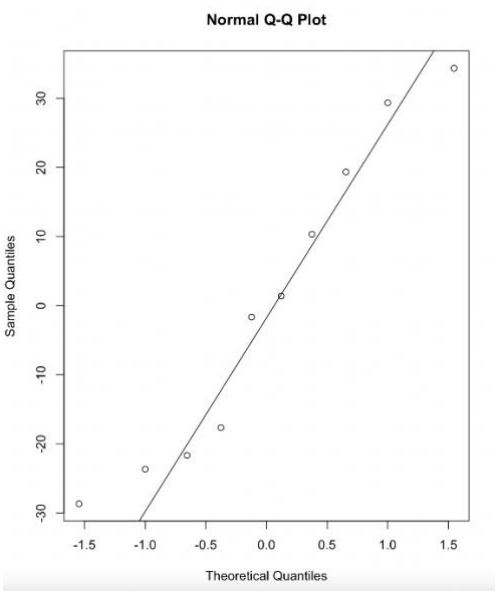
	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	1763.667	1732.644	1794.690	1716.221	1811.112
2022	1797.333	1753.460	1841.206	1730.235	1864.431
2023	1831.000	1777.267	1884.733	1748.822	1913.178
2024	1864.667	1802.621	1926.712	1769.776	1959.557
2025	1898.333	1828.964	1967.703	1792.242	2004.424
2026	1932.000	1856.010	2007.990	1815.783	2048.217
2027	1965.667	1883.588	2047.746	1840.138	2091.195
2028	1999.333	1911.587	2087.079	1865.137	2133.529
2029	2033.000	1939.931	2126.069	1890.664	2175.336
2030	2066.667	1968.564	2164.770	1916.631	2216.702
2031	2100.333	1997.442	2203.225	1942.975	2257.692

R output showing healthcare spending forecast for 2021-2031 with 80% and 95% confidence intervals

The 95% confidence interval around forecasted healthcare spending is tight, whilst the ACF and normal Q-Q plot below indicate a goodness of fit.



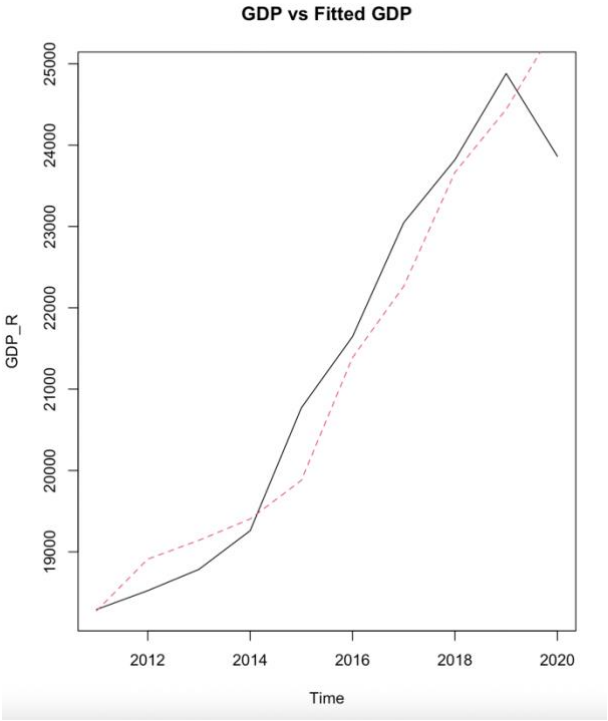
ACF of Fitted Healthcare Residuals



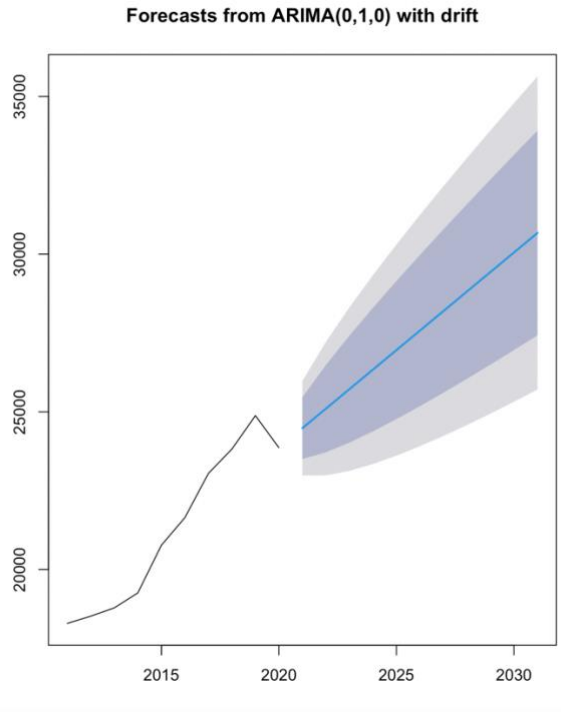
Normal Q-Q Plot of Healthcare Spending Residuals

M. GDP per capita without a national team

GDP per capita was forecasted using through time series using an ARIMA(0,1,0) model with drift. This was calculated in order to understand how the GDP would have continued to grow if there was no implementation of an international soccer team in Rarita.



GDP vs Fitted GDP for 2011-2020 using an ARIMA(0,1,0) model with drift



Forecasted GDP for 2021-2031 using an ARIMA(0,1,0) model with drift

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	24482	23503.35	25460.65	22985.29	25978.71
2022	25101	23716.99	26485.01	22984.33	27217.67
2023	25720	24024.94	27415.06	23127.62	28312.38
2024	26339	24381.71	28296.29	23345.58	29332.42
2025	26958	24769.68	29146.32	23611.25	30304.75
2026	27577	25179.82	29974.18	23910.82	31243.18
2027	28196	25606.75	30785.25	24236.08	32155.92
2028	28815	26046.97	31583.03	24581.66	33048.34
2029	29434	26498.06	32369.94	24943.87	33924.13
2030	30053	26958.25	33147.75	25319.99	34786.01
2031	30672	27426.20	33917.80	25707.97	35636.03

R output showing GDP forecast for 2021-2031 with 80% and 95% confidence intervals

N. GDP per capita with a national team

Data used to predict GDP per capita for 2021-2031 with the impact of an international soccer team
Data from 2011-2020 was used in a linear regression in order to forecast the GDP from 2021-2031. The forecasted outputs as well as the linear regression inputs can be seen in the following graph. The times series forecast for healthcare spending from 2021-2031 was utilised in the model below, as well as the expenses and revenues previously calculated for 2021-2031.

Year	GDP (per capita)	Healthcare Spending (per capita)	Profit (per capita)
2011	18,292	1,427	25
2012	18,523	1,437	26
2013	18,785	1,449	27
2014	19,260	1,465	28
2015	20,770	1,509	29
2016	21,646	1,541	30
2017	23,047	1,604	41
2018	23,820	1,657	22
2019	24,880	1,725	33
2020	23,863	1,730	14
2021	25,187	1,764	17
2022	24,621	1,797	-6
2023	25,807	1,831	2
2024	27,032	1,865	10
2025	28,279	1,898	19
2026	29,529	1,932	28
2027	31,606	1,966	51
2028	33,776	1,999	76
2029	36,195	2,033	106
2030	38,901	2,067	141
2031	41,933	2,100	181

In order to determine the level of profit in Rarita for the years 2011-2015 total revenue and expenses were forecasted backwards using a rate of 0.93 for revenue and 0.92 for expenses. This was calculated by taking the average of *Revenue_t/Revenue_t-1* and the average of *Expenses_t/Expenses_t-1* for the years 2016-2019. The change from 2019/2020 was excluded as it was an outlier.

Economic data for 2020 indicates some sort of market crash due to the stark decrease in GDP, inflation and soccer profit. It was assumed that the years 2011-2015 did not experience such a market crash and thus only the change in revenue and expenses for 2016-2019 was considered.

Year Difference	Rev	Exp
2016/2017	0.9386	1.0112
2017/2018	0.9512	0.8082
2018/2019	0.8906	0.9411
2019/2020	1.1275	1.0129
Average excluding 2019/2020	**0.93**	**0.92**

N. 1. R Code and Output of Linear Regression for GDP

Initially the data for 2011-2020 was partitioned into a training and test data set as seen in the following R code.

```
# partition data into a training and test set
set.seed(2)
dt = sort(sample(nrow(Rarita_completed), nrow(Rarita_completed)*.7))
train<-Rarita_completed[dt,]
test<-Rarita_completed[-dt,]

# fit model using selected variables
fit.train=lm(GDP~Year+Healthcare+Profit,data=train)

pred.train.GDP<-predict(fit.train,newdata = test)
predict(fit.train,newdata=test,interval='confidence')
```

Statistics regarding this model can be seen in the table below

R Squared	Adjusted R Squared	sigma	P value	df	Log likelihood	AIC	BIC
0.991	0.982	324	0.00144	3	-47.4	105	105

The complete data for 2011-2020 was used in order to fit the 2021-2031 data. The summary of this fitted model is detailed below through R output.

Summary of fitted model

```
Call:
lm(formula = GDP ~ Year + Healthcare + Profit, data = Rarita_completed)

Residuals:
    Min      1Q  Median      3Q     Max
-414.31 -247.24  -89.34  195.36  697.24

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -4.451e+05  3.983e+05  -1.118   0.3065
Year         2.185e+02  2.016e+02   1.084   0.3200
Healthcare   1.574e+01  5.214e+00   3.020   0.0234 *
Profit       5.659e+01  2.027e+01   2.792   0.0315 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 417.9 on 6 degrees of freedom
Multiple R-squared:  0.9814,    Adjusted R-squared:  0.9722
F-statistic: 105.7 on 3 and 6 DF,  p-value: 1.39e-05
```

O. Okun's Law

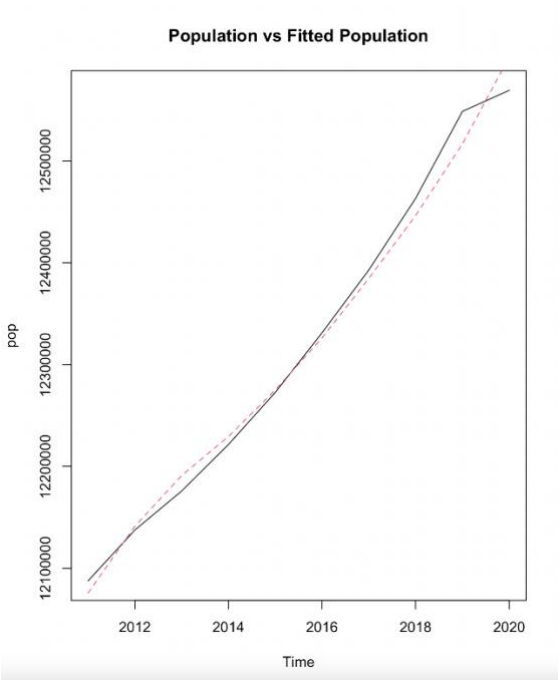
Okun's law, proposed by Arther Okun, investigates the relationship between a countries GDP and unemployment rate.

$$\frac{Y - Y^*}{Y^*} = -c(u - u^*)$$

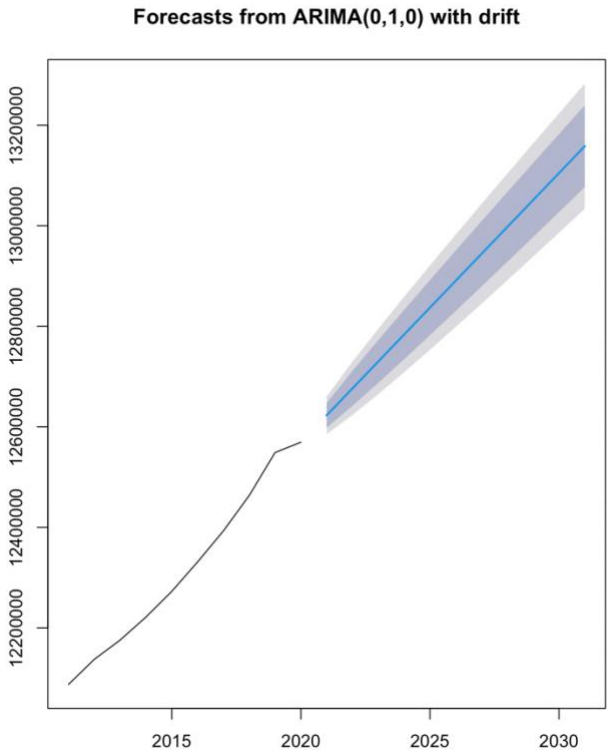
Where;
Y = potential GDP
Y* = actual output
C = factor relating to changes in unemployment to changes in output
U = actual unemployment rate
U* = natural rate of unemployment
Whilst there are other factors at play that influence both GDP and unemployment there seems to be empirical support for Okun's law.

P. Population

The population of Rarita from 2021-2031 was forecast through a time series in R in order to determine the GDP per province in Rarita as detailed below.



Population vs Fitted Population for 2011-2020 using an ARIMA(0,1,0) model with drift



Forecasted Population for 2021-2031 using an ARIMA(0,1,0) model with drift

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	12623010	12598493	12647528	12585514	12660507
2022	12676549	12641875	12711222	12623520	12729577
2023	12730087	12687621	12772554	12665141	12795034
2024	12783626	12734590	12832662	12708632	12858619
2025	12837164	12782341	12891988	12753319	12921010
2026	12890703	12830646	12950759	12798855	12982551
2027	12944241	12879373	13009109	12845034	13043448
2028	12997780	12928433	13067127	12891722	13103837
2029	13051318	12977764	13124872	12938827	13163809
2030	13104856	13027324	13182389	12986281	13223432
2031	13158395	13077078	13239711	13034032	13282758

R output showing Population forecast for 2021-2031 with 80% and 95% confidence intervals

Q. Impact of GDP per province

In order to determine the impact of GDP on each province the population of Rarita from 2021-2031 was projected using a time series. To ascertain how the population would be distributed within Rarita's various provinces, the percentage of population residing in each province was calculated from 2011-2021. The average population percentage from 2011-2021 was assumed to demonstrate how the population would be distributed between the provinces in 2021-2031.

Percentage of Population			
Year	East Rarita	Central Rarita	West Rarita
2011	0.15143	0.25073	0.59784
2012	0.15153	0.24973	0.59874
2013	0.15178	0.24802	0.60020
2014	0.15212	0.24631	0.60157
2015	0.15257	0.24495	0.60248
2016	0.15313	0.24388	0.60299
2017	0.15371	0.24299	0.60330
2018	0.15411	0.24247	0.60342
2019	0.15431	0.24251	0.60318
2020	0.15460	0.24028	0.60512
Average	**0.15293**	**0.245187**	**0.601882878**

Using this average percentage, population per province was calculated as seen in the following table.

Population				
Year	East Rarita	Central Rarita	West Rarita	Rarita
2021	∂ 1,930,437	∂ 3,095,000	∂ 7,597,574	∂ 12,623,010
2022	∂ 1,938,624	∂ 3,108,127	∂ 7,629,798	∂ 12,676,549
2023	∂ 1,946,812	∂ 3,121,254	∂ 7,662,021	∂ 12,730,087
2024	∂ 1,955,000	∂ 3,134,381	∂ 7,694,246	∂ 12,783,626
2025	∂ 1,963,187	∂ 3,147,508	∂ 7,726,469	∂ 12,837,164
2026	∂ 1,971,375	∂ 3,160,635	∂ 7,758,693	∂ 12,890,703
2027	∂ 1,979,562	∂ 3,173,762	∂ 7,790,917	∂ 12,944,241
2028	∂ 1,987,750	∂ 3,186,889	∂ 7,823,141	∂ 12,997,780
2029	∂ 1,995,938	∂ 3,200,015	∂ 7,855,365	∂ 13,051,318
2030	∂ 2,004,125	∂ 3,213,142	∂ 7,887,588	∂ 13,104,856
2031	∂ 2,012,313	∂ 3,226,269	∂ 7,919,813	∂ 13,158,395

To calculate the total GDP of Rarita from 2021-2031 the forecasted population was multiplied with the predicted GDP per capita. The next step was to determine the percentage of GDP attributable to each province from 2021-2031. The percentage of total GDP for each province was calculated for 2011-2020. It can be seen in the table below that East Rarita's percentage of total GDP was increasing at the highest rate, followed by West Rarita, with Central Rarita's percentage of GDP decreasing.

Percentage of GDP			
Year	East Rarita	Central Rarita	West Rarita
2011	0.38181	0.30952	0.30869
2012	0.38624	0.29917	0.31461
2013	0.38912	0.29210	0.31877
2014	0.39411	0.28961	0.31631
2015	0.40697	0.28147	0.31157
2016	0.41155	0.27961	0.30882
2017	0.41379	0.27839	0.30781
2018	0.41022	0.28183	0.30792
2019	0.40343	0.28110	0.31548
2020	0.41161	0.27267	0.31573

As our implementation plan has a focus on decreasing the wealth gap between the provinces, percentage of total GDP per province for 2021-2031 was calculated in accordance. It was assumed that from 2021-2024 the percentage of GDP per province would remain that same as it did in 2020. This is because these years mark the infancy of our newly formed International Team, whilst soccer revenues remained low.

However, from 2025-2031 the percentage of total GDP in Central Rarita was assumed to grow at a rate of 1.5% p.a. and the percentage of total GDP in West Rarita was assumed to grow at a rate of 2.5% p.a.

Percentage of GDP projected for 2021-2031			
Year	East Rarita	Central Rarita	West Rarita
2021	0.41161	0.27267	0.31573
2022	0.41161	0.27267	0.31573
2023	0.41161	0.27267	0.31573
2024	0.41161	0.27267	0.31573
2025	0.39961	0.27676	0.32363
2026	0.38737	0.28091	0.33172
2027	0.37486	0.28513	0.34001
2028	0.36208	0.28940	0.34851
2029	0.34903	0.29375	0.35722
2030	0.33569	0.29815	0.36616
2031	0.32207	0.30262	0.37531

Using these assumptions, we were able to determine the total GDP per province from 2021-2031.

	Total GDP			
Year	East Rarita	Central Rarita	West Rarita	Rarita
2021	ð 130,866,187,059	ð 86,692,858,148	ð 100,384,064,446	ð 317,938,025,012
2022	ð 128,467,349,141	ð 85,103,737,841	ð 98,543,977,976	ð 312,110,073,522
2023	ð 135,224,604,984	ð 89,580,110,504	ð 103,727,294,013	ð 328,526,755,519
2024	ð 142,236,674,132	ð 94,225,285,317	ð 109,106,070,740	ð 345,562,503,763
2025	ð 145,069,062,401	ð 100,472,153,534	ð 117,485,694,572	ð 363,026,910,507
2026	ð 147,450,334,633	ð 106,928,994,139	ð 126,267,790,882	ð 380,647,119,653
2027	ð 153,359,804,451	ð 116,648,841,496	ð 139,102,634,056	ð 409,111,280,004
2028	ð 158,958,871,431	ð 127,051,814,838	ð 153,000,771,277	ð 439,011,457,546
2029	ð 164,879,583,952	ð 138,763,566,250	ð 168,750,870,966	ð 472,394,021,168
2030	ð 171,132,411,108	ð 151,994,231,918	ð 186,661,821,918	ð 509,788,464,945
2031	ð 177,706,618,051	ð 166,978,576,935	ð 207,084,203,542	ð 551,769,398,528

Finally, GDP per capita was calculated as *Total_GDP_t/Population_t* as seen in the table below.

	GDP Per Capita			
Year	East Rarita	Central Rarita	West Rarita	Rarita
2021	ð 67,791	ð 28,011	ð 13,213	ð 25,187
2022	ð 66,267	ð 27,381	ð 12,916	ð 24,621
2023	ð 69,460	ð 28,700	ð 13,538	ð 25,807
2024	ð 72,755	ð 30,062	ð 14,180	ð 27,032
2025	ð 73,895	ð 31,921	ð 15,206	ð 28,279
2026	ð 74,796	ð 33,831	ð 16,274	ð 29,529
2027	ð 77,472	ð 36,754	ð 17,854	ð 31,606
2028	ð 79,969	ð 39,867	ð 19,557	ð 33,776
2029	ð 82,608	ð 43,363	ð 21,482	ð 36,195
2030	ð 85,390	ð 47,304	ð 23,665	ð 38,901
2031	ð 88,310	ð 51,756	ð 26,148	ð 41,933
Growth	30.27%	84.77%	97.90%	66.49%

Bibliography

[i] Goal (November 22, 2019), *"Football loan transfer rules: Premier League guidelines, limits & full details"*, https://www.goal.com/en-om/news/football-loan-transfer-rules-premier-league-guidelines/1kbkbnfluqauo1jpdowcm1pimp

[ii] Megan Garcia (April 4, 2020), *"Positions in Soccer and Their Roles"*, How They Play, https://howtheyplay.com/team-sports/Positions-in-Soccer-and-Their-Roles

[iii] Katherine Lucas (January 20, 2022), *"New loan rules explained, and what Fifa regulations mean for Chelsea, Man City and other Premier League clubs"*, inews, https://inews.co.uk/sport/football/loan-rules-new-explained-fifa-regulations-chelsea-man-city-liverpool-players-1413714

[iv] Jason Brownlee (May 23, 2018), *"A Gentle Introduction to k-fold Cross-Validation"*, Machine Learning Mastery, https://machinelearningmastery.com/k-fold-cross-validation/#:~:text=Cross%2Dvalidation%20is%20a%20resampling,k%2Dfold%20cross%2Dvalidation.