William Davis
STA4224
Final Report
Professor Hon Yiu So

Introduction

In this project, we will analyze the effect of age, height, weight, average yearly salary (APY), and draft pick of National Football League (NFL) players on fantasy football performance. Fantasy football is a game that is played by football fans where players can draft a team, set a lineup every week, and compete against league mates. Fantasy points are scored primarily from offensive NFL players through passing, running, or receiving the football during each game. Different amounts of fantasy points are scored for yards, receptions, touchdowns, etc. A typical fantasy league consists of 12 teams. The goal is to draft the players in each position (quarterback (QB), wide receiver (WR), running back (RB), and tight end (TE)) that score the most fantasy points. Since there are 12 teams in a league, the top 12 scoring players in each position are grouped together. For example, the top 12 RB scorers are referred to as RB1's, the next 12 scorers RB2's, etc. This is the response that we are interested in analyzing. For this analysis, we will focus on the RB and WR positions.

We will be utilizing a cumulative logit model, which provides a latent variable estimate. A latent variable is a variable that is not directly observed, but is assumed through the observed data (Agresti, 2019). We will be able to use the latent variable estimate to predict which category each player will end up in. We will look at some exploratory charts relating our explanatory variables to fantasy finish category for both RB and WR before creating the cumulative logit model. we will also create decision tree models for both RB and WR positions and see if the conclusions are similar to those of the latent response model and to discuss the advantages and drawbacks of each. After conducting the analyses, we will discuss potential improvements that we could incorporate into future analyses on this topic. All Figures, as well as a knitted R Markdown document, can be found in the appendix.

Exploratory Analysis

Before starting the latent response model, we will create exploratory plots to visualize the relationship between our explanatory variablesand our response variable, fantasy finish category. Refer to Figure 1. This plot visualizes the relationship between age, weight, and our response, RB fantasy finish category. As we can see, almost all the RB1s appear to be under the age of 27. So, from this plot, the expectation is that as age increases, the probability of being an RB1 decreases. We see larger chunks of the RB2 and RB3 categories that have ages above 27. We can also see that most RB1s appear to be above 215 pounds. We will compare these visual conclusions with the conclusions of the latent response model later.

Refer to Figure 2. This plot visualizes the relationship between RB fantasy finish category, age, and height. It looks like RBs who are 72 inches or taller are more likely to be RB1s compared to the other categories. We will compare this conclusion with what we find in the latent response model.

Refer to Figure 3. This plot visualizes the relationship between RB fantasy finish category, average yearly salary (in millions of dollars), and the pick number the player was selected with in the NFL draft. From this plot, we would guess that as salary increases, the player is more likely to be in the RB1 category. We would also guess that as draft pick increases (i.e., the player was taken later in the draft), then the player is less likely to finish in the RB1 category. We will keep these visual conclusions in mind when we conduct our running back latent response analysis. Now, we will look at the WR data.

Refer to Figure 4. This plot visualizes the relationship between WR fantasy finish category, age, and weight. There is not a single WR1 over the age of 30. In general, it looks like there is not much correlation between fantasy finish category and age until the age of 29 or 30, when a drop-off occurs. So, based on the visual, we would expect the probability of being a WR1 or WR2 to decrease as age increases. Then, it is difficult to see any clear relationship with weight and WR category.

Refer to Figure 5. This plot visualizes the relationship between WR fantasy finish category, age, and height. It is difficult to spot a clear trend with this plot. All categories look evenly spread across heights. We will be able to draw a conclusion from the output of the latent response model.

Refer to Figure 6. This plot visualizes the relationship between WR fantasy finish category, average yearly salary (millions), and draft pick number. Similar to RBs, it appears that as salary increases, the player is more likely to finish in the WR1 category. Also similar to RBs, as draft pick increases, the player appears less likely to finish in the WR1 category. We will be able to compare these conclusions to the model.

Statistical Method

The method that will be used for this analysis is the cumulative logit model for ordinal responses. The model can be represented by $P(Y \leq j) = \pi_1 + \pi_2 + \cdots + \pi_j, j = 1, \ldots, c$. The probabilities are consistent with the ordering of the responses, ie. $P(Y \leq 1) \leq P(Y \leq 2) \leq \cdots \leq P(Y \leq c) = 1$. The cumulative logits are defined as $\text{logit}\big(P(y \leq j)\big) = \log\left(\frac{P(y \leq j)}{1 - P(Y \leq j)}\right) = \log\left(\frac{\pi_1 + \cdots + \pi_j}{\pi_{j+1} + \cdots + \pi_c}\right), j = 1, \ldots, c-1$ (Agresti, 2019). This model implies a latent variable, which is estimated by $Y^* = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \epsilon$, where ϵ has mean 0 and constant variance at all values of independent variables. Then, the ordinal categorical variable cumulative probabilities can be represented by $\text{logit}[P(Y \leq j)] = \alpha_j - \beta_1 x_1 - \beta_2 x_2 - \cdots - \beta_p x_p,\ j = 1, \ldots, c-1$ (Agresti, 2019). We could also replace the logit function with other linking functions, such as probit (the inverse of the normal CDF) or tanh (the hyperbolic tangent function). In this study, we will focus on the logit linking function. Now that we have the foundation, we will take a look at the analysis.

Statistical Analysis

For running backs, our coefficients are as follows:

age 0.1376820
height -0.1400196
weight -0.0212671
APY -0.1535811
pick -0.0008904

Our intercepts are as follows:

RB1|RB2 -13.0165
RB2|RB3 -11.7001
RB3|RB4 -10.4443

The coefficient on age is positive, which implies that as a player gets older, they are less likely to be an RB1. Then, the coefficients on height, weight, APY, and pick are negative, which implies that taller/heavier players, player with higher salaries, and players that were picked later are more likely to be RB1s. Outside of the pick coefficient, these are consistent with our visual conclusions. Note that the coefficient on pick is very close to 0. We are likely dealing with multicollinearity, as we would not expect a negative coefficient on pick. We can get a given RBs latent variable estimate from the equation discussed above. For example, take the Detroit Lions RB, D'Andre Swift. Swift is 69 inches tall, weighs 211 pounds, was 22.9863 years old as of the end of the 2021 NFL season, has an average yearly salary of 2.13473 million, and was selected with the 35th pick of the draft. We can find Swift's latent variable estimate by $Y^* = 0.137682 * 22.9863 - 0.1400196 * 69 - 0.0212671 * 211 - 0.1535111 * 2.13473 - 0.0008904 * 35 = -11.3429$. So, based on D'Andre Swift's age, height, and weight alone we would predict that he was an RB2 since $-11.7001 < -11.3429 < -13.0165$. By calculating the latent variable estimate for all RBs, we can then arrange them in ascending order to classify each as RB1, RB2, RB3, or RB4. The smaller the latent variable estimate, the more likely the RB is to be an RB1. Refer to Figures 7-10. These tables include player name, category, and latent rank. These tables give us an idea of how our model did in predicting categories RB1-4. To expand the information conveyed by our model, we can run a simulation that will create epsilons following a standard logistic regression and adding the epsilon to the latent variable estimate in each stage of the simulation. By doing so, we can determine probabilities that each player is in each category by adding up the number of occurrences in each category after the simulation and dividing by 1000. This allows for a deeper interpretation of the model. For the RB position, we can see the summary tables that include probabilities for each category in Figures 11-14. we separated each Figure by the predicted category before the simulation was ran. The simulation can be done in R using a for loop. To simulate 1000 error terms, we can use the rlogis() function in R, which will generate random numbers following a standard logistic regression. In each simulation, we will add an error term to each player's latent variable estimate. The code for the simulation can be found in the knitted R markdown portion of the Appendix.

For wide receivers, the coefficients are as follows:

age 0.172443
height 0.229591
weight -0.015963
APY -0.065534
pick 0.005957

and the intercepts are as follows:

WR1|WR2	16.7929
WR2|WR3	18.0802
WR3|WR4	19.4449

The coefficients on age, height, and pick are positive, which implies that older/taller WRs and WRs that were selected later in the draft are less likely to be WR1s. The coefficients on weight and APY are negative, which implies that heavier WRs and WRs with higher average yearly salaries are more likely to be WR1s. We are also likely experiencing multicollinearity in this model, as intuitively we would not expect taller WRs to be less likely to finish in the WR1 category. A given player's latent variable could be calculated just as we did for D'Andre Swift with the RB estimates. Then, we can calculate the latent variable estimates for all WRs to classify each as predicted to be WR1, WR2, WR3, or WR4. Refer to Figures 15-18. These tables give us an idea on how our model did in predicting each category of WR. Like the RB position, we can run a simulation of error terms to determine the probability of each category for each player. After running 1000 simulations, we divided the total number of occurrences that each WR landed in each category and divided by 1000 to get the probability. The summary tables that contain the probabilities can be seen in Figures 19-22, which are separated by which category the WRs were predicted to be in before adding error terms. Displaying these probabilities allow someone who is not familiar with statistical methods to have a better understanding of the results.

Decision Tree Model

In a post written by Kailash Awati, the rpart function in R is described to "…split the data recursively, which means that the subjects that arrive from a split are further split until a predetermined termination criterion is reached." The goal of the algorithm is to increase homogeneity in each node after each split. It is important to note that the tree splits in the ideal way for the given dataset, which leads to decision trees being at risk for overfitting to the given data (Awati, 2016). A decision tree could appear to predict perfectly but would not do well when introduced to new data. Because of this, it is preferred to split that data into a training set and a test set to see how the tree does on new data.

Using the rpart function in R, we created a decision tree based on the same data set. The objective is to compare the conclusions of the decision tree model to see if they are similar to that of the cumulative logit model. Refer to Figure 23. This is the decision tree for the RB position. We can see that if weight is greater than 220 pounds and the player was selected before pick number 70, then RB1 is predicted. This is consistent with the conclusions on weight of the cumulative logit model in that heavier RBs were more likely to be RB1s. However, the cumulative logit model has a negative coefficient on pick, which would imply that later picks are more likely to be RB1s. The tree model disagrees and concludes that earlier draft picks are more likely to be RB1s than later draft picks. This is more aligned with what our intuition would tell us. Refer to Figure 24, the decision tree for the WR position. The WR tree model causes some concerns for overfitting. The only explanatory variable used is pick. The tree model says if the player was picked at 175 or later, then RB4 is predicted. Then, if the player was picked earlier than 175 and later than 60, RB1 is predicted. If the player was picked earlier than 175, and earlier than 44, then RB1 is predicted. If the player was selected earlier than 175, earlier than 61, and later than 44, RB3 is predicted. Since there is only one season worth of data, it is plausible that the decision tree is overfit to the data. We would not expect these arbitrary draft pick cutoff points to work well with new data. Another downside of our WR decision tree model is that it does not predict any WR2s. This is a weakness in comparison to the cumulative logit model, which can give us 12 players predicted to be in each category. One potential advantage, however, of decision tree models is that it they are much easier to explain to someone who is not well-versed with data and statistics compared to the cumulative logit model.

Limitations

There are a few limitations to these models. One limitation for the cumulative logit model is that it does not account for the constraint that each response category consists of only 12 players. A limitation of the decision tree model is that it does not account for the ordinality of the response categories. Another limitation of the decision tree model is that since we have not yet used it on test data, we do not know how well it would do in practice. Our WR model appears to be using arbitrary cutoff points that would not work well when introduced to new data. Ideally, we would test the tree models on a test set before using them in practice.

Opportunities for Further Analysis

It is natural to think of ways to improve an analysis while it is in progress or once it is completed. There are a few things that we would like to add to this analysis in the future. One possible method for improvement would be to include several seasons of data instead of just one. we could create data frames for each season and create the age column based on the players age at the end of the given season and then join all the seasons together. A larger sample could improve the accuracy/usefulness of the model. Then, as the same players were measured over several seasons, we could consider a mixed effect model for the cumulative regression model. The random effect could then be viewed as a skill effect for each player. Another potential opportunity for improvement would be to add more independent variables. One change to the independent variables to consider is instead of using average salary per year for a player’s entire contract, to use average guaranteed money per year. NFL contracts are often created with in-game incentives that are only paid if the incentive is met. One player who has a large contract, mostly due to incentives, is WR Christian Kirk. Kirk had the lowest latent variable estimate of all WRs but was a WR3 this year. Perhaps it would be better to use guaranteed money per year instead of including incentives. This is something that we would like to investigate in the future.

Conclusions

In this study, we utilized a cumulative logit model for ordinal responses and a decision tree model to model fantasy finish category for the running back and wide receiver positions. Using the cumulative logit model, we were able to use the latent variable estimate to predict which category each player landed in. Then, we were able to run 1000 simulations on the latent response variable where we added the error term that follows a standard logistic regression. By doing so, we were able to find the probability that each player lands in each response category. This allows for easier interpretation of the results. Then, a decision tree model was used to compare to the conclusions of the two models. There are some concerns for overfitting with the tree models, particularly the WR tree. It is interesting to evaluate the differences between the two models. Furthermore, we discussed potential opportunities for improvement such as increasing sample size, utilizing a mixed effects model, and introducing more independent variables.

References

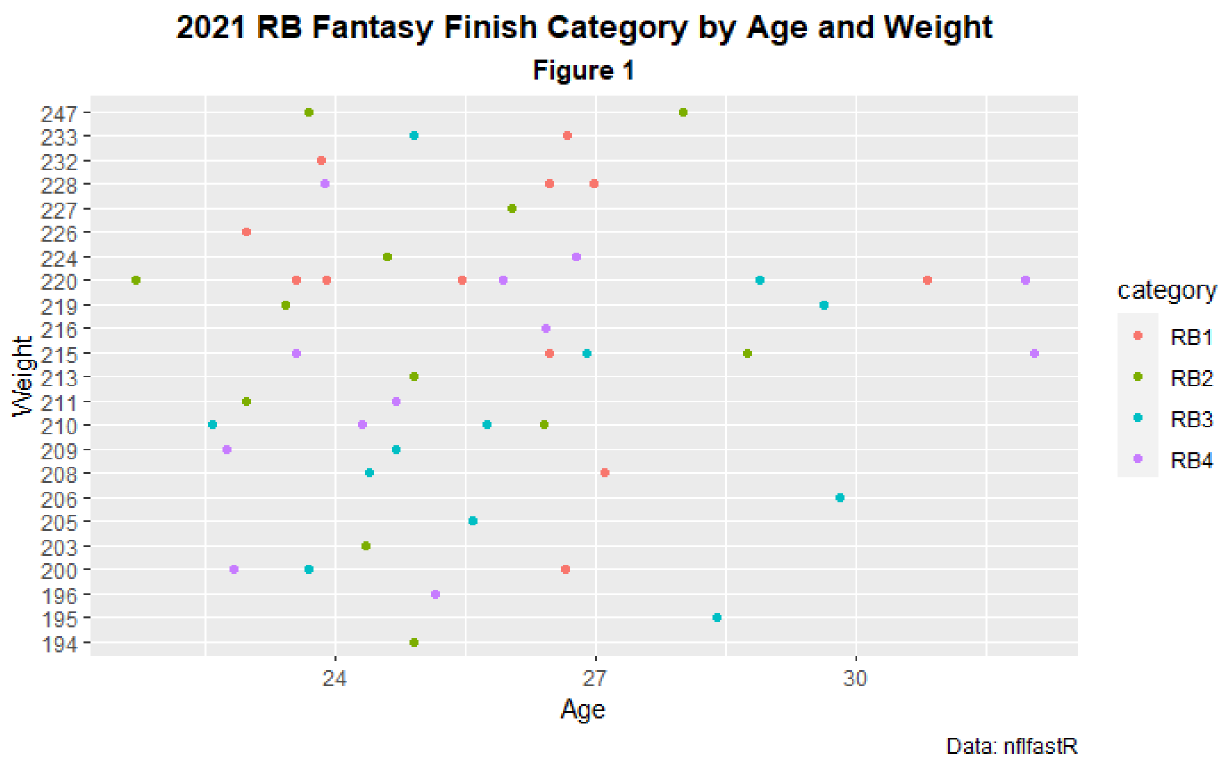
Agresti, A. (2019). An Introduction to Categorical Data Analysis. John Wiley & Sons.

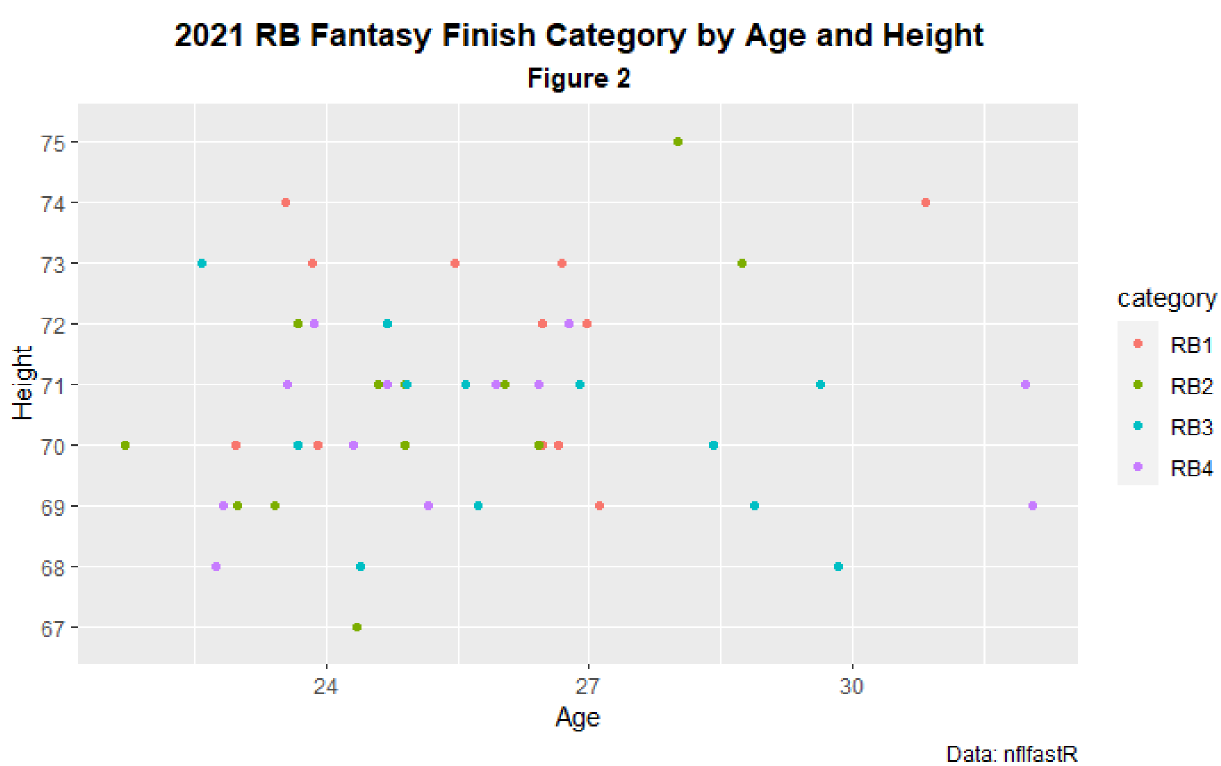
Awati, K. (2016). A gentle introduction to decision trees using R. https://eight2late.wordpress.com/2016/02/16/a-gentle-introduction-to-decisiontrees-usingr/#:~:text=The%20rpart%20algorithm%20works%20by,predetermined%20termin ation%20criterion%20is%20reached.

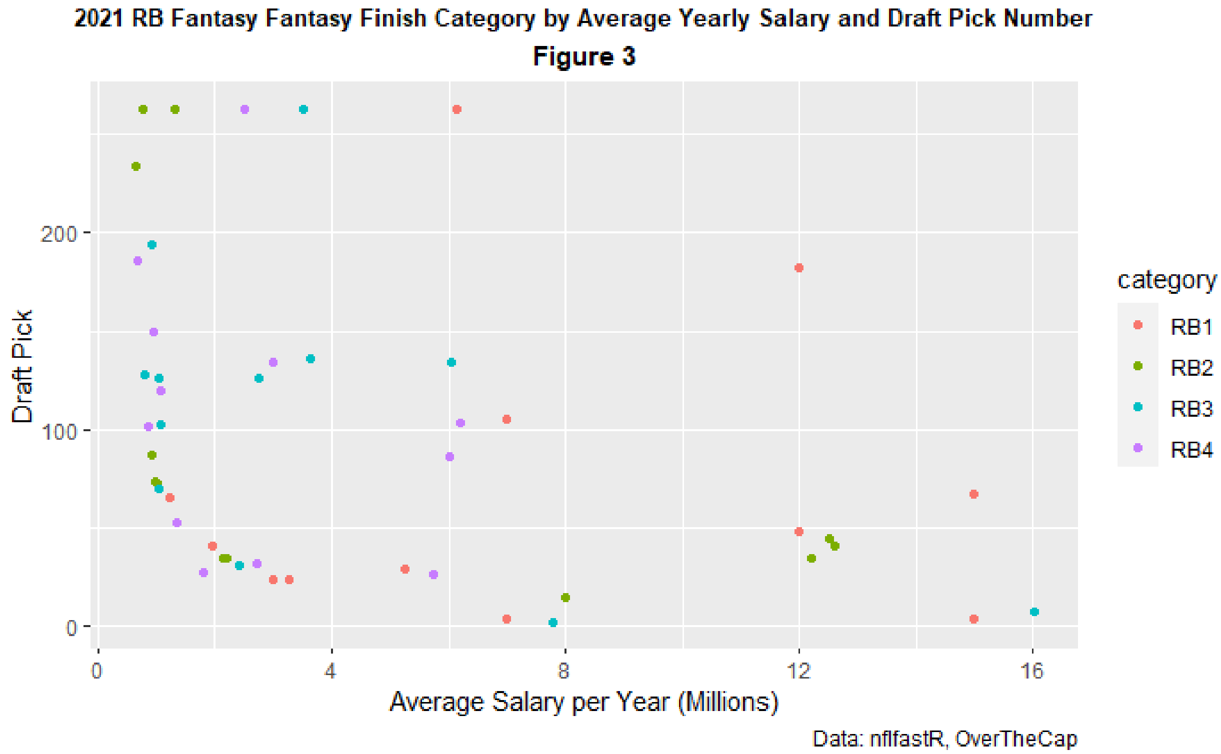
https://www.nflfastr.com/

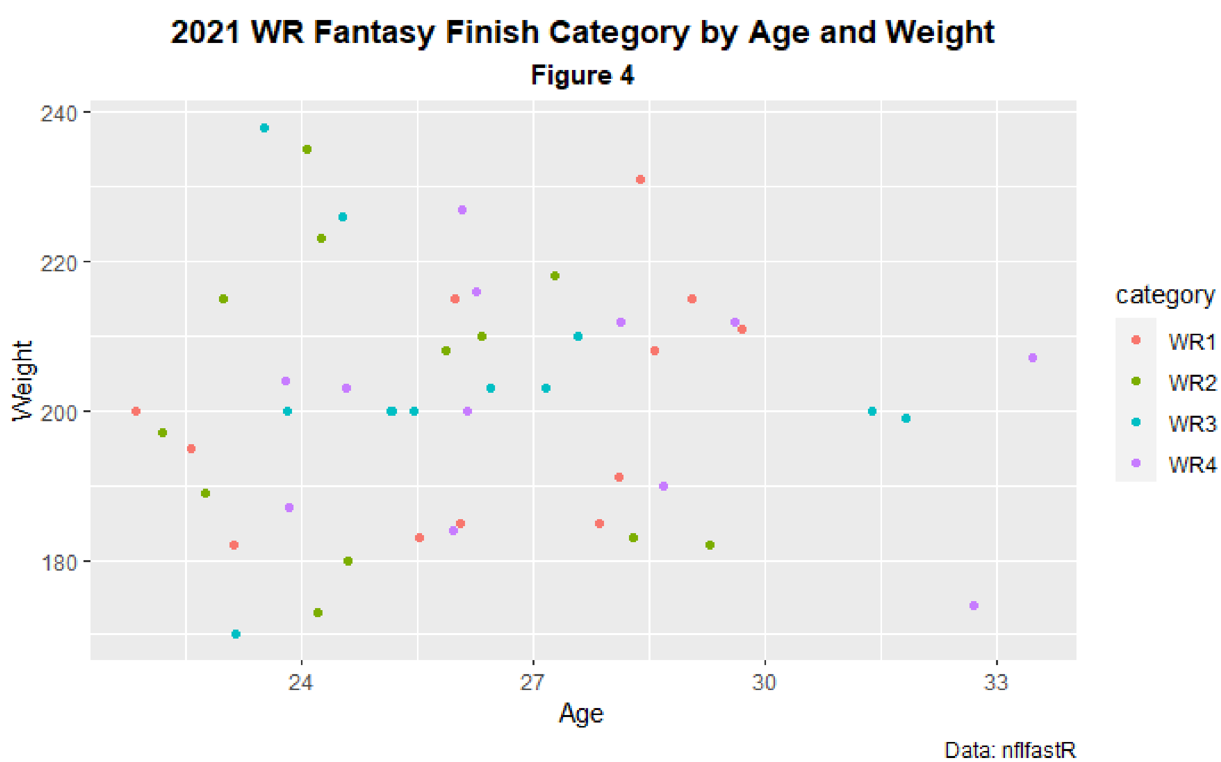
https://overthecap.com/contracts/

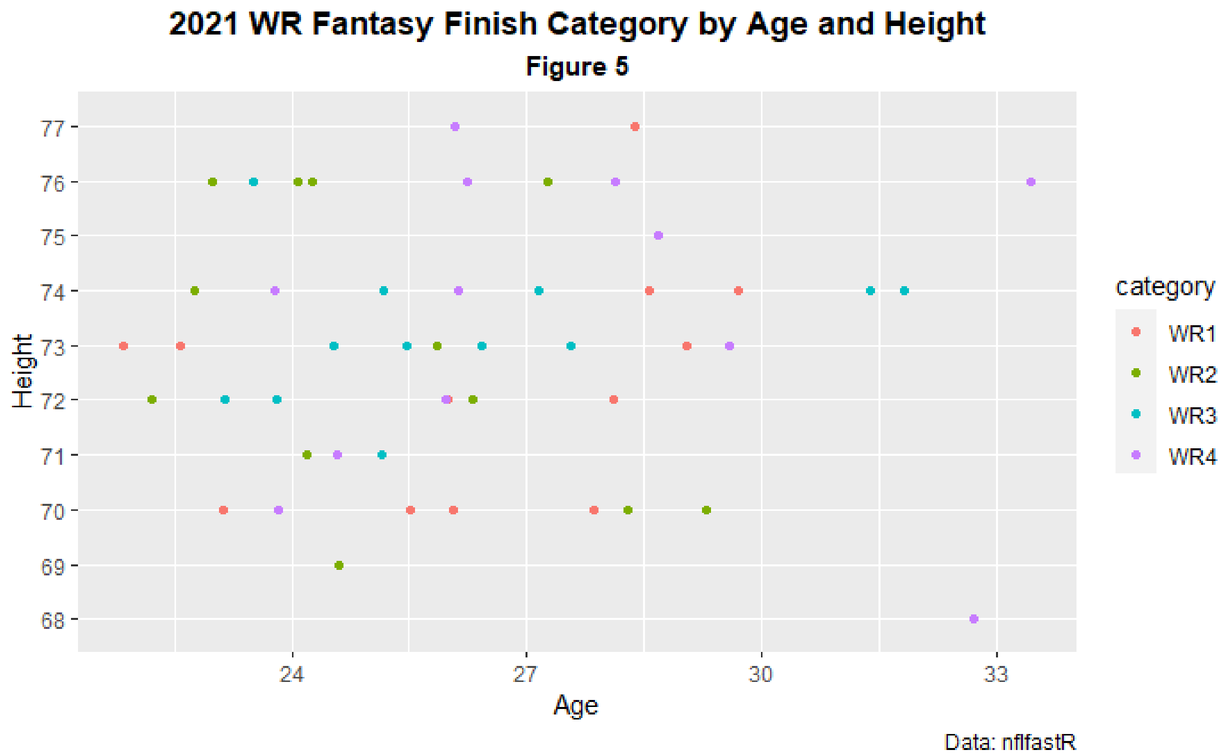
Appendix

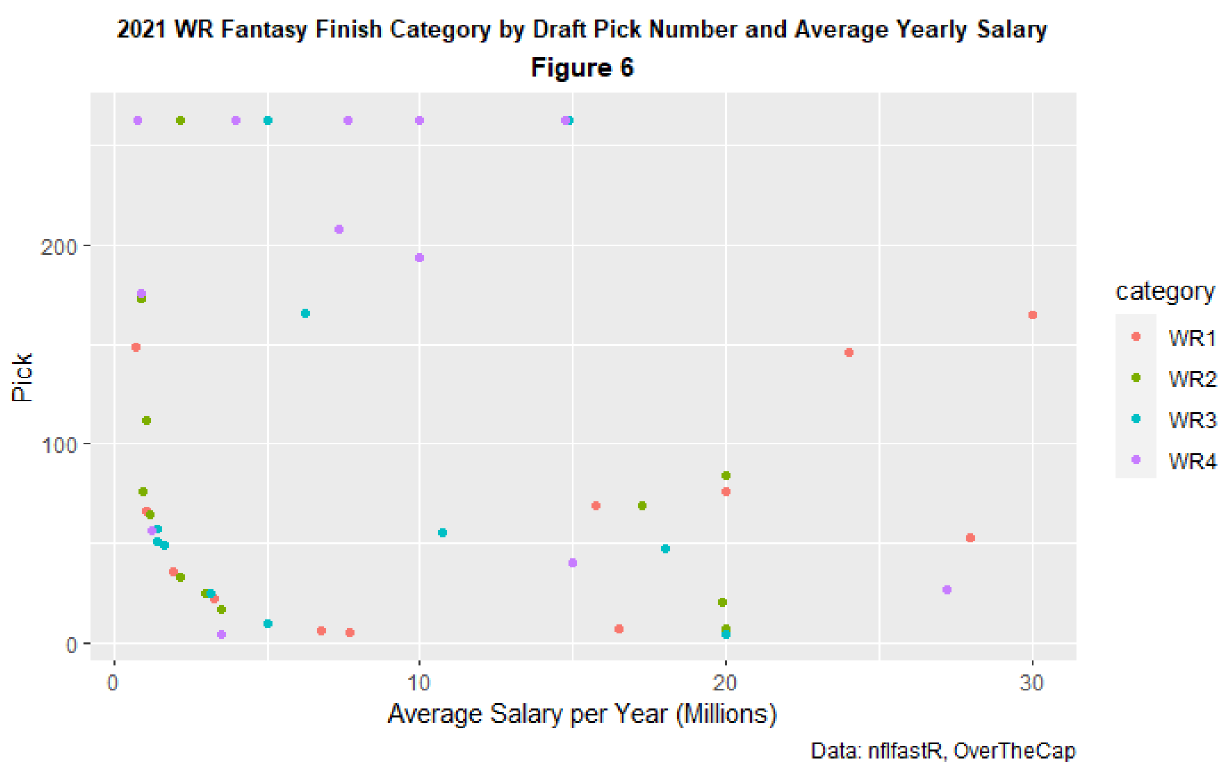

2021 RB Fantasy Finish Category by Age and Weight
Figure 1
Weight
247
233
232
228
227
226
224
220
219
216
215
213
211
210
209
208
206
205
203
200
196
195
194
Age
24
27
30
category
RB1
RB2
RB3
RB4
Data: nflfastR


2021 RB Fantasy Finish Category by Age and Height
Figure 2
Height
75
74
73
72
71
70
69
68
67
Age
24
27
30
category
RB1
RB2
RB3
RB4
Data: nflfastR


2021 RB Fantasy Fantasy Finish Category by Average Yearly Salary and Draft Pick Number
Figure 3
Draft Pick
200
100
0
0
4
8
12
16
Average Salary per Year (Millions)
category
RB1
RB2
RB3
RB4
Data: nflfastR, OverTheCap


2021 WR Fantasy Finish Category by Age and Weight
Figure 4
Weight
240
220
200
180
24
27
30
33
Age
category
WR1
WR2
WR3
WR4
Data: nflfastR


2021 WR Fantasy Finish Category by Age and Height
Figure 5
Height
Age
category
WR1
WR2
WR3
WR4
Data: nflfastR


2021 WR Fantasy Finish Category by Draft Pick Number and Average Yearly Salary
Figure 6
Pick
Average Salary per Year (Millions)
category
WR1
WR2
WR3
WR4
Data: nflfastR, OverTheCap

Predicted RB1s

Figure 7

player_name1	category	latent_rank
D.Henry	RB2	1
E.Elliott	RB1	2
J.Mixon	RB1	3
C.McCaffrey	RB3	4
A.Kamara	RB1	5
N.Chubb	RB2	6
J.Conner	RB1	7
S.Barkley	RB3	8
D.Cook	RB2	9
A.Dillon	RB2	10
N.Harris	RB1	11
A.Jones	RB1	12

Predicted RB2s

Figure 8

player_name1	category	latent_rank
L.Fournette	RB1	13
M.Gordon	RB2	14
A.Gibson	RB1	15
R.Penny	RB4	16
R.Stevenson	RB4	17
K.Hunt	RB4	18
J.Williams	RB2	19
C.Hubbard	RB3	20
J.Taylor	RB1	21
J.Williams	RB4	22
J.Jacobs	RB1	23
C.Edmonds	RB3	24

Predicted RB3s

Figure 9

player_name1	category	latent_rank
C.Patterson	RB1	25
A.Ekeler	RB1	26
D.Montgomery	RB2	27
A.Mattison	RB4	28
J.Robinson	RB2	29
N.Hines	RB4	30
T.Pollard	RB3	31
D.Swift	RB2	32
M.Sanders	RB4	33
C.Edwards-Helaire	RB4	34
D.Harris	RB2	35
S.Michel	RB3	36

Predicted RB4s

Figure 10

player_name1	category	latent_rank
D.Booker	RB3	37
T.Johnson	RB4	38
E.Mitchell	RB3	39
K.Gainwell	RB4	40
M.Davis	RB3	41
B.Bolden	RB4	42
D.Henderson	RB3	43
J.McKissic	RB3	44
M.Gaskin	RB2	45
D.Singletary	RB2	46
M.Ingram	RB4	47
D.Freeman	RB3	48

RB Category Probabilities for Predicted RB1s

Figure 11

player_name1	category	latent_rank	RB1_probability	RB2_probability	RB3_probability	RB4_probability
D.Henry	RB2	1	0.613	0.179	0.130	0.078
E.Elliott	RB1	2	0.576	0.208	0.130	0.086
J.Mixon	RB1	3	0.576	0.206	0.146	0.072
C.McCaffrey	RB3	4	0.524	0.245	0.146	0.085
A.Kamara	RB1	5	0.522	0.230	0.152	0.096
N.Chubb	RB2	6	0.501	0.254	0.149	0.096
J.Conner	RB1	7	0.470	0.257	0.177	0.096
S.Barkley	RB3	8	0.452	0.257	0.176	0.115
D.Cook	RB2	9	0.419	0.267	0.188	0.126
A.Dillon	RB2	10	0.391	0.283	0.200	0.126
N.Harris	RB1	11	0.364	0.313	0.209	0.114
A.Jones	RB1	12	0.375	0.280	0.213	0.132

RB Category Probabilities for Predicted RB2s

Figure 12

player_name1	category	latent_rank	RB1_probability	RB2_probability	RB3_probability	RB4_probability
L.Fournette	RB1	13	0.377	0.271	0.215	0.137
M.Gordon	RB2	14	0.346	0.281	0.235	0.138
A.Gibson	RB1	15	0.311	0.301	0.221	0.167
R.Penny	RB4	16	0.303	0.278	0.242	0.177
R.Stevenson	RB4	17	0.270	0.303	0.258	0.169
K.Hunt	RB4	18	0.294	0.281	0.253	0.172
J.Williams	RB2	19	0.297	0.273	0.233	0.197
C.Hubbard	RB3	20	0.238	0.288	0.274	0.200
J.Taylor	RB1	21	0.225	0.309	0.261	0.205
J.Williams	RB4	22	0.221	0.309	0.274	0.196
J.Jacobs	RB1	23	0.233	0.284	0.272	0.211
C.Edmonds	RB3	24	0.214	0.279	0.288	0.219

RB Category Probabilities for Predicted RB3s

Figure 13

player_name1	category	latent_rank	RB1_probability	RB2_probability	RB3_probability	RB4_probability
C.Patterson	RB1	25	0.204	0.268	0.296	0.232
A.Ekeler	RB1	26	0.187	0.292	0.274	0.247
D.Montgomery	RB2	27	0.185	0.269	0.272	0.274
A.Mattison	RB4	28	0.164	0.291	0.310	0.235
J.Robinson	RB2	29	0.165	0.276	0.310	0.249
N.Hines	RB4	30	0.153	0.282	0.285	0.280
T.Pollard	RB3	31	0.150	0.288	0.297	0.265
D.Swift	RB2	32	0.135	0.259	0.305	0.301
M.Sanders	RB4	33	0.141	0.272	0.290	0.297
C.Edwards-Helaire	RB4	34	0.128	0.252	0.306	0.314
D.Harris	RB2	35	0.137	0.256	0.277	0.330
S.Michel	RB3	36	0.132	0.223	0.295	0.350

RB Category Probabilities for Predicted RB4s

Figure 14

player_name1	category	latent_rank	RB1_probability	RB2_probability	RB3_probability	RB4_probability
D.Booker	RB3	37	0.110	0.215	0.318	0.357
T.Johnson	RB4	38	0.092	0.242	0.303	0.363
E.Mitchell	RB3	39	0.105	0.224	0.297	0.374
K.Gainwell	RB4	40	0.102	0.218	0.292	0.388
M.Davis	RB3	41	0.089	0.194	0.285	0.432
B.Bolden	RB4	42	0.080	0.186	0.299	0.435
D.Henderson	RB3	43	0.076	0.178	0.307	0.439
J.McKissic	RB3	44	0.085	0.194	0.260	0.461
M.Gaskin	RB2	45	0.064	0.182	0.293	0.461
D.Singletary	RB2	46	0.072	0.193	0.252	0.483
M.Ingram	RB4	47	0.073	0.153	0.284	0.490
D.Freeman	RB3	48	0.059	0.157	0.251	0.533

Predicted WR1s

Figure 15

player_name	category	latent_rank
C.Kirk	WR3	1
J.Waddle	WR1	2
D.Adams	WR1	3
B.Cooks	WR2	4
D.Hopkins	WR4	5
J.Chase	WR1	6
A.Cooper	WR3	7
T.Hill	WR1	8
C.Godwin	WR2	9
M.Brown	WR2	10
B.Aiyuk	WR3	11
M.Williams	WR2	12

Predicted WR2s

Figure 16

player_name	category	latent_rank
M.Hardman	WR4	13
J.Jefferson	WR1	14
T.Lockett	WR2	15
D.Smith	WR3	16
A.Brown	WR3	17
S.Diggs	WR1	18
D.Samuel	WR1	19
C.Lamb	WR2	20
C.Sutton	WR4	21
A.St. Brown	WR2	22
M.Evans	WR1	23
D.Johnson	WR1	24

Predicted WR3s

Figure 17

player_name	category	latent_rank
K.Allen	WR1	25
C.Claypool	WR3	26
C.Kupp	WR1	27
T.Higgins	WR2	28
T.Boyd	WR3	29
T.McLaurin	WR2	30
D.Metcalf	WR2	31
V.Jefferson	WR3	32
K.Osborn	WR4	33
H.Renfrow	WR1	34
R.Gage	WR4	35
D.Mooney	WR2	36

Predicted WR4s

Figure 18

player_name	category	latent_rank
C.Wilson	WR4	37
K.Bourne	WR3	38
M.Callaway	WR4	39
J.Meyers	WR3	40
M.Pittman	WR2	41
C.Beasley	WR4	42
A.Green	WR4	43
R.Anderson	WR4	44
A.Thielen	WR3	45
T.Patrick	WR4	46
A.Lazard	WR4	47
M.Jones	WR3	48

WR Category Probabilities for Predicted WR1s

Figure 19

player_name1	category	latent_rank	WR1_probability	WR2_probability	WR3_probability	WR4_probability
C.Kirk	WR3	1	0.543	0.277	0.147	0.033
J.Waddle	WR1	2	0.502	0.274	0.189	0.035
D.Adams	WR1	3	0.539	0.260	0.164	0.037
B.Cooks	WR2	4	0.478	0.296	0.179	0.047
D.Hopkins	WR4	5	0.465	0.284	0.205	0.046
J.Chase	WR1	6	0.472	0.275	0.199	0.054
A.Cooper	WR3	7	0.444	0.291	0.211	0.054
T.Hill	WR1	8	0.427	0.288	0.230	0.055
C.Godwin	WR2	9	0.430	0.316	0.197	0.057
M.Brown	WR2	10	0.378	0.309	0.232	0.081
B.Aiyuk	WR3	11	0.371	0.304	0.233	0.092
M.Williams	WR2	12	0.378	0.290	0.252	0.080

WR Category Probabilities for Predicted WR2s

Figure 20

player_name1	category	latent_rank	WR1_probability	WR2_probability	WR3_probability	WR4_probability
M.Hardman	WR4	13	0.366	0.313	0.234	0.087
J.Jefferson	WR1	14	0.376	0.289	0.244	0.091
T.Lockett	WR2	15	0.364	0.286	0.255	0.095
D.Smith	WR3	16	0.339	0.308	0.250	0.103
A.Brown	WR3	17	0.315	0.330	0.237	0.118
S.Diggs	WR1	18	0.312	0.296	0.272	0.120
D.Samuel	WR1	19	0.279	0.336	0.248	0.137
C.Lamb	WR2	20	0.278	0.311	0.285	0.126
C.Sutton	WR4	21	0.269	0.312	0.276	0.143
A.St. Brown	WR2	22	0.218	0.331	0.287	0.164
M.Evans	WR1	23	0.229	0.292	0.320	0.159
D.Johnson	WR1	24	0.213	0.314	0.310	0.163

WR Category Probabilities for Predicted WR3s

Figure 21

player_name1	category	latent_rank	WR1_probability	WR2_probability	WR3_probability	WR4_probability
K.Allen	WR1	25	0.225	0.263	0.329	0.183
C.Claypool	WR3	26	0.207	0.278	0.324	0.191
C.Kupp	WR1	27	0.218	0.292	0.259	0.231
T.Higgins	WR2	28	0.177	0.267	0.332	0.224
T.Boyd	WR3	29	0.173	0.276	0.320	0.231
T.McLaurin	WR2	30	0.168	0.235	0.345	0.252
D.Metcalf	WR2	31	0.151	0.262	0.312	0.275
V.Jefferson	WR3	32	0.152	0.236	0.315	0.297
K.Osborn	WR4	33	0.144	0.230	0.283	0.343
H.Renfrow	WR1	34	0.133	0.252	0.266	0.349
R.Gage	WR4	35	0.126	0.244	0.276	0.354
D.Mooney	WR2	36	0.128	0.217	0.272	0.383

WR Category Probabilities for Predicted WR4s

Figure 22

player_name1	category	latent_rank	WR1_probability	WR2_probability	WR3_probability	WR4_probability
C.Wilson	WR4	37	0.117	0.203	0.261	0.419
K.Bourne	WR3	38	0.120	0.183	0.273	0.424
M.Callaway	WR4	39	0.119	0.173	0.252	0.456
J.Meyers	WR3	40	0.095	0.157	0.256	0.492
M.Pittman	WR2	41	0.091	0.166	0.233	0.510
C.Beasley	WR4	42	0.077	0.153	0.228	0.542
A.Green	WR4	43	0.076	0.142	0.243	0.539
R.Anderson	WR4	44	0.076	0.119	0.230	0.575
A.Thielen	WR3	45	0.085	0.120	0.185	0.610
T.Patrick	WR4	46	0.054	0.128	0.182	0.636
A.Lazard	WR4	47	0.050	0.114	0.195	0.641
M.Jones	WR3	48	0.053	0.108	0.173	0.666

Figure 23

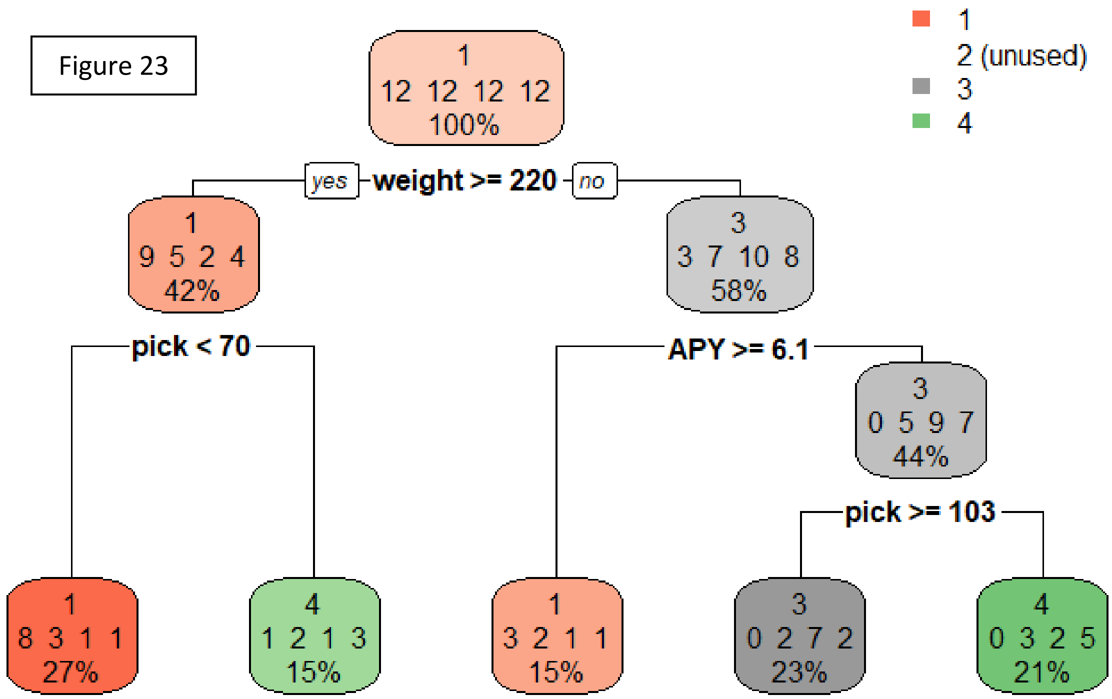


Figure 24