

Predicting MLB Likelihood Using Vizual Edge Scores

Alex Kikos

Introduction

Every year, all 30 Major League Baseball (MLB) teams meet in the Summer for the amateur draft, hoping to find the future star(s) of their franchise. Thousands of high school and collegiate amateur players hope to hear their name called by the MLB Commissioner as they look to make their mark on professional baseball. Unlike other pro sport leagues, the MLB has *several* minor-league levels (Rookie, A-, A, A+, AA, AAA) that nearly all players must progress through in order to reach the mountaintop of baseball, the MLB. In baseball, position-players (non-pitchers) are evaluated mainly on hitting, power, running, fielding and throwing, which has been dubbed the ‘Five Tools of Baseball’. MLB clubs hire their own scouts to evaluate a prospect’s “five tools”. Yet, one of the most obvious aspects about a hitter’s ability is their vision. After all, MLB hitters have ~0.40s to react to an incoming pitch and just as the old saying in baseball goes, “*you can’t hit what you can’t see*”. While every player is required to go through a standard physical, including an eye exam, those eye exams are intended to assess whether a player needs glasses/contacts and does not indicate if a player can accurately perceive depth or can track an object. Insert Vizual Edge.

Vizual Edge is a leading visual skills assessment and training company that uses 3D-technology with its web-based application, the Edge Trainer, to assess an individual’s visual processing skills (how the eyes and brain work *together*). MLB teams hire Vizual Edge so they can administer the Vizual Edge test to potential draft picks and add a *sixth* dimension to the common “Five Tools” that make up a player. The 10-minute test measures a player’s Eye Alignment, Depth Perception, Convergence, Divergence, Recognition and Tracking. A composite score (“Edge Score”) is scored out of 100 and takes all 6 of the exercises into account, providing scouts and coaches with a benchmark number for a player’s overall visual ability. Though Vizual Edge has been around for over 20 years and collected 20,000+ scores from MLB draft prospects alone, the company has never taken an in-depth look at its own database to find potential correlations with players reaching the MLB. Simple analyses have been conducted to find general trends athletes, and usually better players tend to have higher Edge Scores (see Figure 1), but advanced data analysis has yet to be run. The purpose of this study was to **provide a verified study on the predictive power of the Vizual Edge test scores and MLB-likelihood**. Not only will this provide Vizual Edge immense value by justifying an increase in licensing fees for use of the program, but MLB teams can also better understand *how* to interpret strictly-vision data to help identify draft prospects who could make it to the highest level in pro baseball (MLB), rather than being stuck in the minor leagues. Additionally, this analysis can help the MLB teams save millions of dollars of unnecessary contracts to players who are least likely to make it to the MLB, which at the end of the day, is of the utmost importance to the organizations.

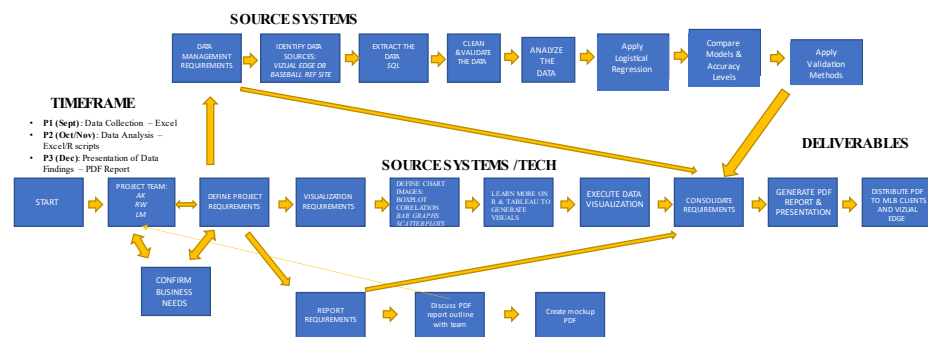
Figure 1: Breakdown of average Vizual Edge scores from 1,233 players at every level in baseball

	Pro Level	# of Players	Edge Score	Alignment	Depth Perception	Convergence	Divergence	Recognition Accuracy	Recognition Speed (s)	Tracking Accuracy	Tracking Speed (s)
MLB	299	83.32	1	94%	47	27	98%	0.81	96%	0.49	
AAA	217	81.92	1	90%	44	27	97%	0.90	95%	0.51	
AA	167	81.08	1	88%	41	25	97%	0.88	95%	0.50	
A+	222	80.79	1	88%	41	24	96%	0.90	94%	0.51	
A	151	79.98	1	85%	39	23	96%	0.99	94%	0.52	
A-	80	78.90	1	82%	37	23	95%	0.96	93%	0.53	
RK	97	79.51	1	84%	38	22	94%	1.03	94%	0.55	

BI Methods

Prior to beginning this study, a general framework needed to be built on what the end goal was, what was required, how do we get there and what methods should be used. After speaking with my project team at Vizual Edge, CEO (RW) & Director of Baseball (LM), we decided to look at how each of the individual Vizual Edge test scores, along with the composite 'Edge Score', influenced a player's likelihood of reaching Major League Baseball. It should be stated that Vizual Edge never claims that their test scores are the sole reasoning of an athlete's ability (or inability), given that numerous external factors influence talent. That being said, the purpose of this study was to look strictly at the Vizual Edge data, and not include external factors such as swing or biomechanical data, which are not typically publicly available. Due to strict privacy protocols imposed by MLB teams, inquiring about desired outcomes directly with our MLB clients was not feasible. However, our Director of Baseball worked as an MLB team front office executive for over 15+ years, which helped gain a better understanding of what MLB clubs are interested in seeing as it pertains to new, verified data. The project team decided that since the Vizual Edge data was stored internally, there was no financial commitment required to jumpstart this study. Figure 2 above outlines the overall business plan with requirements needed to execute an analysis of this magnitude. The project team expected to see the following outcomes from the study: **summary of Vizual Edge data** per level in baseball, **probability of reaching the MLB**, **hypothesis testing** involving Edge Scores at/above 83.0 and the **most suitable logistic regression model** for predicting MLB likelihood. The project team decided that it would be best to focus on the **2010-2015 MLB draft years**, as this provided ample time for players who were drafted to potentially reach the MLB by now (2021). Cross-referencing draft results during those years and position-players who took at least one Vizual Edge test, it was determined that **1,233 players** would form the data set used for this entire analysis. The Vizual Edge data was stored on a server, that only a select few individuals have access to, including myself. After receiving permission to utilize from the Vizual Edge CEO, athlete data was anonymized by creating a unique ID (UID). In addition to the Vizual Edge data (which included the 7 variables listed above), a few publicly available metrics were collected for each player as well, using BaseballReference.com. These additional metrics included a combination of numerical (Draft Year, Draft Round, Draft Pick and Draft Age) and categorical data (Level Reached, Draft Type and Status). All data was stored in an Excel spreadsheet and later imported into R, a statistical programming language.

Figure 2: An overview of the study's deliverables and process to reach end goal.



Since it is common for players to take the Vizual Edge test multiple teams (usually due to multiple teams being interested in drafting the player), several unique rows of data exist for a single player. After importing the data set into R, additional data management needed to be done to the multiple scores per player. Upon the advice of our Director of Baseball, the best Vizual Edge scores for each player were extracted and placed into a new data frame.

Alternative methods of taking a player's *average* scores were considered, however, it is widely accepted in baseball to take a player's peak scores to represent their current ceiling. It is important to note that for three Vizual Edge variables (Alignment, Recognition2 and Tracking2), the 'best' score refers to the lowest value scored. After initial data management was done, categorical values needed to be converted into numeric ones. This was a required

Figure 3: Sample rows of Vizual Edge and dummy variables

1	UID	Edge	Depth	Convergence	Divergence	Recognition1	Tracking1	DraftYear	DraftAge	DraftRound	DraftPick
2	1	66.9	1.00	1	2	1.000	0.990	2012	20	16	515
3	2	85.1	1.00	64	13	0.972	0.962	2013	22	21	629
4	3	86.0	1.00	44	34	1.000	1.000	2012	22	29	892
5	4	86.5	1.00	62	26	0.982	1.000	2014	19	6	179
6	5	78.8	0.75	13	40	1.000	0.938	2014	22	40	1186
7	6	78.9	1.00	34	8	1.000	0.938	2012	17	10	313

1	UID	Alignment	Recognition2	Tracking2	LevelReached	DraftType	Status	ReachedMLB	CollegePlayer	StillActive	LevelReachedNum
2	1	0	0.64	0.51	A	College	Released	0	1	0	3
3	2	0	0.82	0.46	A	College	Released	0	1	0	3
4	3	2	0.76	0.51	A	College	Released	0	1	0	3
5	4	0	1.03	0.50	A	HS	Released	0	0	0	3
6	5	1	1.08	0.49	A	College	Released	0	1	0	3
7	6	0	0.54	0.49	A	HS	Active	0	0	1	3

step to run any sort of statistical testing and allowed for simpler execution of visuals. The variables "DraftType" and "Status" were converted into 'dummy' binary values, assigning a value of 1 to "College" values in

"DraftType" and "Active" for "Status," with all other values turning into zeros. Two data transformations were done with the "LevelReached" data frame: assigning a variable to each level and turning MLB vs non-MLB into binary values. Newly-created variable "LevelReachedNum" represented the seven different levels in pro baseball that a player might reach, starting with RK (1) up through the MLB (7). Samples of the data set used can be seen above in Figure 3. A multitude of statistical methods were deployed to ensure the outcomes specified by the project team were achieved. These methods included **probability, hypothesis testing, confidence intervals and logistic regression**. Before diving into advanced analyses, it was important to summarize all the baseline data for the 1,233 players. Vizual Edge averages for each level in pro baseball were generated for a quick overview of how scores differ as players make it to higher levels. A sample of Edge Scores by level can be found in

Figure 4: All 1,233 player Edge Scores and their highest level reached on numeric scale, 7 being MLB.

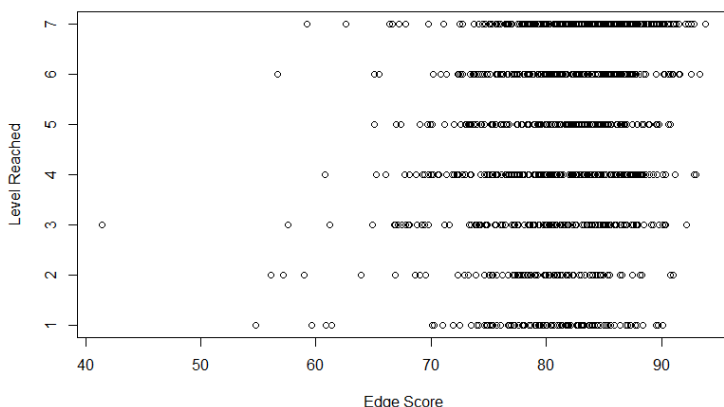


Figure 6 (Results), with further outputs found in Figure 1 and Appendix A1. The primary score of focus was the overall Edge Score, however, as shown in Figure 4, there was not a clear linear relationship between Edge Scores and Level Reached. Thus, identifying outliers was critical to limit the negative influence on summary statistics, such as the standard deviation, mean and inter-quartile range. A Boxplot was created to visually identify outliers that existed in this dataset, of which

31 were discovered. The Boxplot output and summary can be found in the Results section,

with Figure 9. Probability was an extremely useful and applicable tool for this study, based on what we were trying to achieve with predicting MLB likelihood. Initially, Empirical Probability was deployed to assess the likelihood of the Vizual Edge variables to influence reaching the MLB. For example, out of the 1,233 players in this data set, only 299 reached the MLB, thus the probability of reaching the MLB was 0.24. More advanced probability tests, including Conditional Probability, were conducted and their findings are summarized in the Results section, with Table 1. Statistical inference was also utilized in this study, including confidence intervals and hypothesis testing. Confidence intervals at the 95%

Figure 5: Hypothesis Test for MLB Edge Score

$H_0: \mu < 83.0$

$H_a: \mu \geq 83.0$

confidence level were generated for each of the Vizual Edge scores. An important part of this study was hypothesis testing. Without knowing the *true* average Edge Score for every single MLB player, we wanted to examine and test if the average was in fact at or above 83.0, which is approximately what the sample mean was for the MLB data set. Thus, the hypotheses that were created can be found in Figure 5.

Finally, regression was critical to achieving the outcomes outlined by the project team. Without employing regression, it would have been extremely difficult to assess the predictive power of the Vizual Edge data, combined with draft data. Linear regression was initially utilized, and while it certainly had its applications for this study, it was not required for the exact outcome we were looking for. Using a linear model would have focused on predicting the *exact* level a player reached, rather than if they'll make the MLB or not. Future analyses can certainly benefit and would be encouraged to utilize linear regression for predicting what level a player will reach just by looking at their scores. Models were still created to show how one might utilize and interpret the data by incorporating a linear regression model. Without diving too deeply into the results of the linear regression, some models produced an adjusted R^2 value of 0.20, suggesting that 20% of the sample variation could be explained by Vizual Edge data. Similarly, other regression models were considered, such as quadratic and logarithmic, but were ultimately not used due to the slope not fitting with any of these non-linear models. Ultimately, logistic regression models were used to predict whether (yes or no) a player reached the MLB, which was the initial goal of this study and analysis. Several models were generated and compared in order to find the most accurate predictor of reaching the MLB, all of which producing nearly 80% accuracy levels. Unlike linear regression, which relies on goodness-of-fit measures like the standard error, coefficient determination and adjusted R^2 to compare models, logistic regression relies on alternative methods to compare models. Thus, cross-validation methods were deployed to measure model accuracy levels. The k-fold cross-validation method was determined to be more effective than using the Holdout method, as it was less sensitive to data partitioning. The accuracy levels of the three models all hovered approximately around the 80% mark, suggesting that the models were roughly equal in their predictive power. It is important to note that these models and accuracy levels do not reflect the predictive power of solely measuring players who *actually* made it to the MLB, rather, it encompasses both sets of players; those who made it (true positives) and those who didn't (true negatives). To further assess the exact predictiveness of Vizual Edge data only on players who truly made it to the MLB (true positives), one would need to deploy sensitivity metrics.

Results

This study looked at 1,233 pro baseball players drafted and signed between 2010-2015 who completed at least one Vizual Edge test. Figure 6 summarizes the number of players whose highest level ever reached was one of the seven possible pro levels through the 2021 season. The Edge Score is a metric developed by Vizual Edge to provide players, coaches and scouts with a benchmark number that assesses a player's overall visual rating.

Understanding what the Edge Score profile of each level in professional baseball can provide scouts and MLB clubs with a better idea of the general criteria it takes for to reach that level. Figure 7 on the right shows the average Edge Score by each level in pro baseball. This should theoretically come as no surprise, as the best baseball players in the world (those at the MLB level) *should* have the best overall scores. The Edge Score profiles are almost perfectly linear, suggesting that better players (those playing at a higher level) have better vision (higher Edge Scores). Prior to getting into regression and more predictive analysis, it was important to visualize any possible correlations between data. The correlation plot shown in Figure 8 provides an overview of all correlations between variables used in this study. From this figure, we can assume there is some level of correlation between Edge, Depth, Recognition2 and DraftPick with the ReachedMLB variable. It was important to identify any outliers in the data set, which required the creation of a boxplot. From the boxplot, which can be found in Figure 9, 31

Figure 6: Frequency of players who reached their highest level of the seven possible pro baseball levels

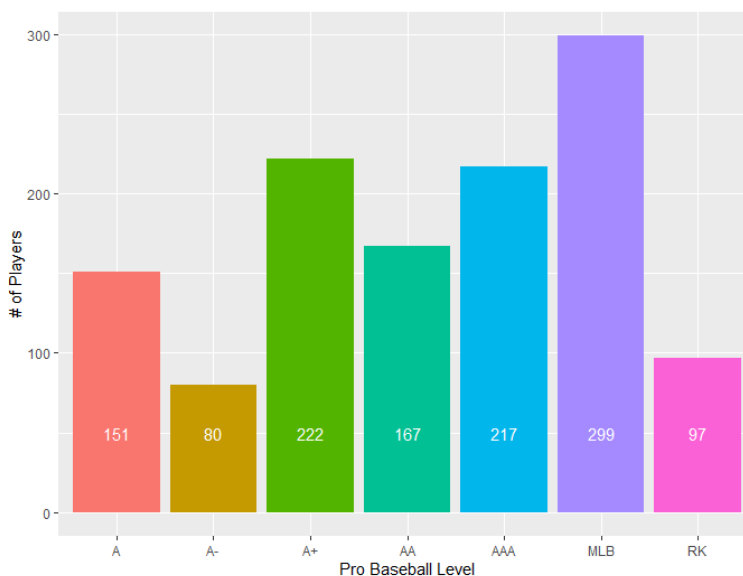


Figure 7: Edge Score Profiles by Pro Level

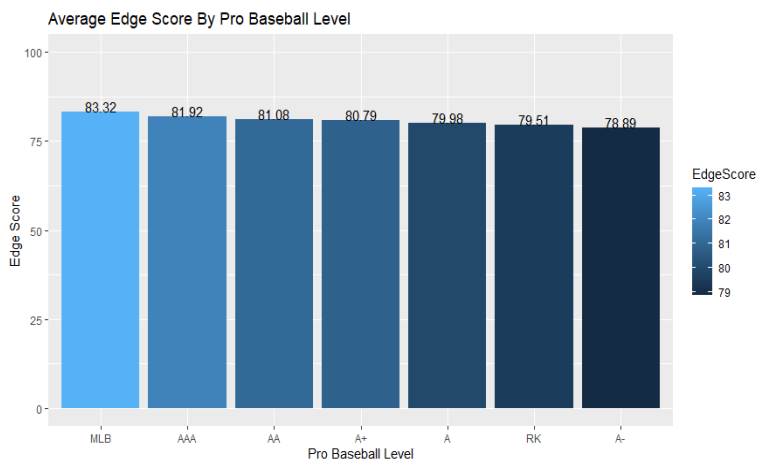


Figure 8: Vizual Edge data set correlation plot

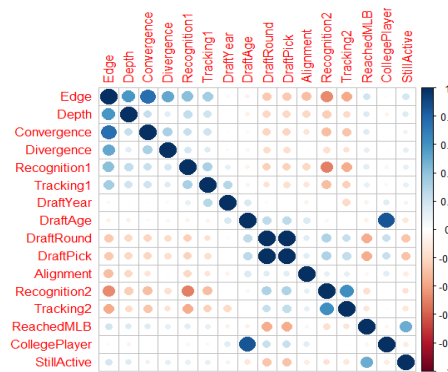
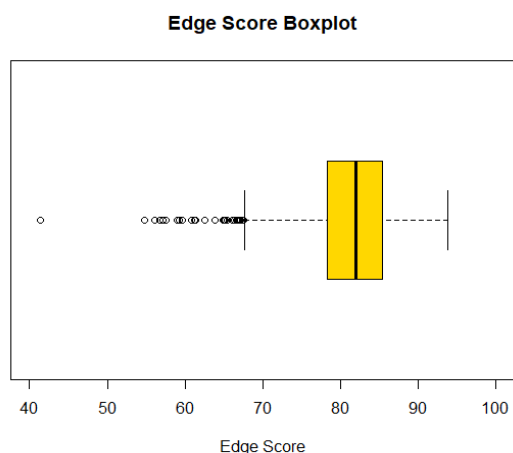


Figure 9: Edge Score Boxplot



arguably the most important visual skill for baseball players, as it relates to a hitter's ability to focus on incoming pitches. The higher the Convergence score, the better the player is at focusing on a pitch. Thus, it is no surprise that it is the lowest value here since it is a critical skill to possess and not every hitter is an *elite* hitter. Additionally, based on the values provided in Table 1, we can attempt to see what the probability is a player reaches the MLB, given they scored at/above 83.0 Edge. Using Conditional Probability, we apply the formula $P(A/B) = P(A \cap B) / P(B)$, where P(A) is 'Reach the MLB' and P(B) is 'Edge Score ≥ 83.0 '. Based on these values, we can conclude that the probability of these events happening is 0.33, or 33%.

After conducting probability tests, more statistical inference techniques were deployed, which included hypothesis testing and confidence intervals. Leveraging statistical packages in R and the mean (81.32) and standard deviation (5.87) of the Edge Scores, a two-tailed test was conducted. With 95% confidence, it can be concluded that the average Edge Score of all pro baseball players (RK through MLB) is between 80.99 and 81.65. A hypothesis test was created to test whether the average Edge Score of MLB players was at/above 83.0. Figure 5 above shows that the null hypothesis in this case is " < 83.0 " and the alternate hypothesis is " ≥ 83.0 ". A t-test was then generated to find the p-value of the hypothesis, which was found to be 0.9963, significantly larger than α (0.05). Thus, at the 5% significance level, we cannot conclude the average MLB Edge Score differs from the average of 83.32, which means we do not reject the null hypothesis.

Edge Scores (~2.5%) were determined to be outliers, all of which were to the left of the IQR. In addition, the Median value for Edge was to the *right* of the Mean, suggesting the data is negatively skewed, or left skewed. Probability was utilized in a variety of ways for this study. Table 1 below summarizes the findings and empirical probabilities for each of the Vizual Edge scores based on the averages calculated for the MLB level. From this table, we can conclude that is fairly likely (68%) that a player in this data set will score at or above 94% on the Depth portion of the test, compared to just a 42% chance scoring above a 47 score on Convergence, the lowest probability value from the MLB averages. Convergence is

	# Of Observations	Probability
Edge Score ≥ 83.0	525	0.43
Depth ≥ 0.94 (94%)	839	0.68
Convergence ≥ 47	522	0.42
Divergence ≥ 27	545	0.44
Recognition2 $\leq 0.81s$	571	0.46
Tracking2 $\leq 0.49s$	566	0.46
Reach The MLB	299	0.24
Reach MLB & Edge ≥ 83.0	171	0.14

Table 1: Probabilities based on MLB average scores

Linear regression was considered during the analysis process, however, given that the data set is not linear by nature (see Figure 4), it was not feasible to utilize. Additionally, considering the intended outcome of this study was to see whether or not (yes or no) a player reached the MLB, linear regression would not be a suitable model to predict this outcome, as better regression options exist. It is important to note that future studies certainly can utilize linear regression to predict the exact level a player will reach, based on the same variables used in this study.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Standard Error of Estimate s_e	1.6890	1.6880	1.6990	1.6890	1.6900	1.6900
Adjusted R^2	0.2049	0.2051	0.1951	0.2046	0.2039	0.2035

Table 2: Goodness-of-fit measures for the six linear regression models.

Six linear regression models were created, as shown above in Table 2, which can be utilized for future studies looking to do further analysis on the exact level a player will reach. Note that the Coefficient of Determination R^2 were not included in Table 2, as the six models had different numbers of predictor variables. Model 2 was found to be the best model out of the six, as a result of having the highest Adjusted R^2 value. This suggests that approximately 20% of sample variation in ‘Level Reached’ can be explained by Model 2. Given that only visual scores from Vizual Edge were used, and zero hitting, swinging or any baseball-related metrics were used, this was quite a high value to find.

Logistic regression was ultimately the most applicable and thus was determined to be the best regression analysis option to proceed with. Since the goal was to determine whether a player made it to the MLB strictly by Vizual Edge scores, logistic regression helped identify that. Figure 10 below outlines the exact three logistic models compared with the highest accuracy levels. All of three of the models consistently generated overall accuracy levels of ~80%. The models were also trained and tested using a K-Fold cross validation method, also with accuracy levels of ~80%.

Figure 10: Top 3 Linear Regression Models

Model 1:

$$\text{ReachedMLB} = -5.8114 + 0.0555E + 0.3889D - 0.0062C - 0.0030DV + 0.0725DA - 0.0046DP$$

Model 2:

$$\text{ReachedMLB} = -6.3772 + 0.0686E - 0.0081C - 0.0045DV + 0.0713DA - 0.0046DP$$

Model 3:

$$\text{ReachedMLB} = -1.1632 + 1.0789D + 0.0042C + 0.3244CP - 0.0046DP$$

Where E = Edge Score, D = Depth, C = Convergence, DV = Divergence, DA = Draft Age, DP = Draft Pick and CP = College Player

Conclusion

The goal of this analysis was to put the Vizual Edge baseline scoring to the test to see if there truly was any predictability level in players reaching the MLB. Having sat on thousands of data points from working closely with MLB teams for a number of years, nothing like this had been done to give MLB teams (and Vizual Edge as a whole) the confidence level of the validation of the Vizual Edge testing scores. Vizual Edge is the first to admit that their scores do not make up the *entire* athlete and that there are a wide variety of

characteristics that make a big-leaguer, such as their size, swing, fielding and more. However, by providing teams with an extra edge in identifying talent in the annual amateur draft, the value on that information could be considered priceless. Simple values found during the analysis process, such as the Edge Score profiles by leagues, is very useful for teams to know that players at those levels are, and for Vizual Edge to incorporate into marketing campaigns and contract negotiations. Even though the logistic models used only generated overall accuracy levels of approximately 80%, future studies may aim to look at the sensitivity rate and other statistical measurements to identify the players that *actually* made it the MLB, rather than the accuracy level of those who did and did not make it. The three logistic regression models that were selected all provided roughly 80% accuracy levels strictly on the Vizual Edge data and any of the three could be utilized today to help predict if a player will reach the MLB.

MLB clubs have entire departments dedicated to finding metrics on a player to try and gain a leg up on their counterparts each year in the draft, and no club looks solely at just one metric when making a decision. That being said, the age of data is quickly accelerating clubs to make more informed decisions based on the information available to them, rather than strictly going on a more subjective measurement of a player. This provides an excellent opportunity for companies like Vizual Edge to have a validated study to provide their MLB clubs confidence that their scores and data are a reliable datapoint on their own, but can be even more vital when combined with the baseball-data collected by scouting departments. Without having access to some of the baseball metrics teams are able to collect on a player, the Vizual Edge scores can only provide so much when evaluating a player, which is why the Adjusted R^2 values in Table 1 were so low. Future studies certainly should attempt to collect as many datapoints on a player as possible, in conjunction with the Vizual Edge scores, in order to improve model accuracy levels and predictiveness. A future study that can predict the exact level a draftee is going to reach, would be incredibly beneficial for MLB teams so they know what a player's ceiling is. This will drastically help contract values, saving teams millions of dollars on players less likely to reach that level. This is a huge problem in the MLB, as large percentage of players drafted do not ever reach the MLB and are often stuck in the minor league system until they burn out and retire. Partnering with companies who collect more baseball-related data might be a business opportunity for both Vizual Edge and said companies to further validate the Vizual Edge test with swing-related metrics, for example.

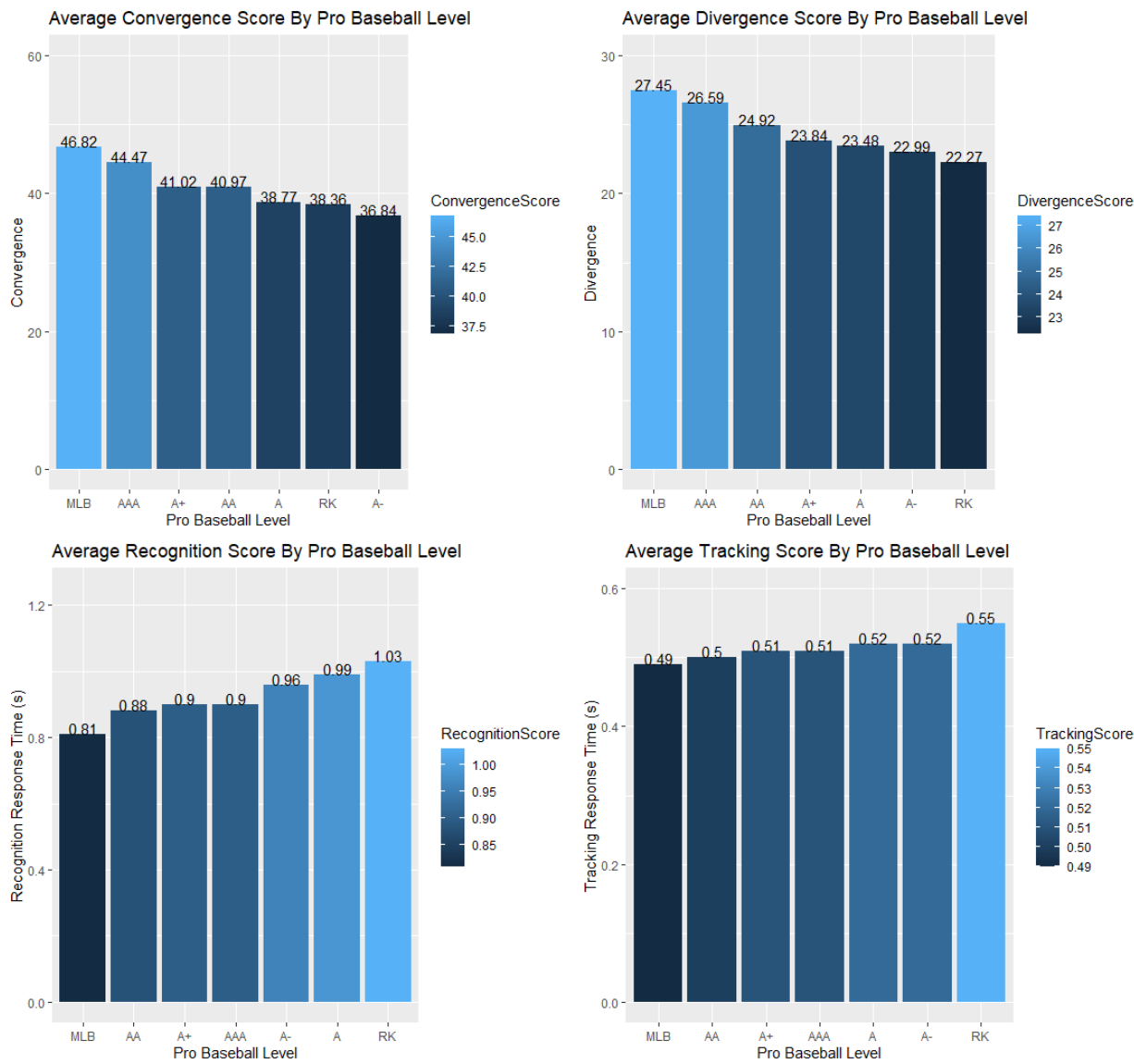
One of the many goals of Vizual Edge is helping and preparing young athletes for what professional teams look for when evaluating talent. That's why Vizual Edge works with so many young baseball players to help realize the importance of the Vizual Edge scores. From the player's point of view, knowing what Vizual Edge scores it takes to reach the MLB can be beneficial by not only training to improve scores prior to meeting with a MLB club to take the test, but also for a personal examination to see if they have what it takes to reach the ultimate level in baseball.

At the end of the day, there is no exact formula that can predict a baseball player reaching the MLB, and the best anyone (team or company) can hope for is to get as close to a perfect formula or model as possible. That was the goal of this analysis, and will continue to be the goal of all future analyses conducted regarding Vizual Edge data. There are an infinite number of variables that make up professional baseball players, and the results extracted from this study confirm what Vizual Edge has marketed for years – *better players have better*

visual scores. Putting a confidence behind our scoring will help with contract negotiations between Vizual Edge and current MLB clients and should help smooth over any concerns about the validity of our data with skeptical, perspective clients. The baseball world is often a tough one to integrate with without having validated data and analyses conducted, which is why a study like this was vital for the future success of Vizual Edge in baseball and MLB contracts. More work certainly can and will be done to provide even more validation of the Vizual Edge scores/test, but this analysis is certainly an excellent starting point to give Vizual Edge the credibility it deserves with its visual skills test.

APPENDIX

A1 - Average Convergence, Divergence, Recognition & Tracking scores by each level in pro baseball



A2: Full summary of Linear Regression models and goodness-of-fit measurements.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	4.1967* (0.0098)	2.9718* (0.0204)	1.0263 (0.4749)	4.9469* (0.0000)	3.3576* (0.0071)	2.1055* (0.0033)
Edge	NA	NA	NA	NA	0.0386* (0.0000)	0.0393* (0.0000)
Depth	0.5889* (0.0265)	0.6005* (0.0236)	0.6055* (0.0234)	0.66945* (0.0106)	NA	NA
Convergence	0.0038 (0.1959)	0.0038 (0.1955)	0.0040 (0.1827)	0.0041 (0.1671)	NA	NA
Divergence	0.0071 (0.1075)	0.0073 (0.0984)	0.0077 (0.0794)	0.0080 (0.0667)	NA	NA
Recognition1	1.3044 (0.1889)	1.3031 (0.1894)	1.2414 (0.2141)	NA	NA	NA
Recognition2	-0.0063 (0.9766)	-0.0151 (0.9444)	-0.0451 (0.8353)	-0.1405 (0.4859)	NA	NA
Tracking1	0.6709 (0.4450)	0.6823 (0.4373)	0.6646 (0.4522)	NA	NA	NA
Tracking2	-0.9449 (0.2414)	-0.9238 (0.2520)	-0.9913 (0.2220)	-1.0058 (0.2117)	NA	NA
DraftAge	-0.0671 (0.2189)	NA	0.1226* (0.0000)	NA	- 0.06670 (0.2203)	NA
CollegePlayer	- 0.0025* (0.0000)	0.6138* (0.0000)	NA	-0.6071* (0.0000)	0.8247* (0.0000)	0.6176* (0.0000)
DraftPick	0.8211* (0.0000)	- 0.0025* (0.0000)	- 0.0024* (0.0000)	-0.0025* (0.0000)	- 0.0026* (0.0000)	- 0.0026* (0.0000)
Standard Error of Estimate s_e	1.6890	1.6880	1.6990	1.6890	1.6900	1.6900
Adjusted R^2	0.2049	0.2051	0.1951	0.2046	0.2039	0.2035