

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

GLEYSONN MENANDRO DE SOUZA

**SPORTS ANALYTICS: VARIÁVEIS SIGNIFICANTES NA SELEÇÃO DE
QUARTERBACKS NA NFL**

UBERLÂNDIA-MG

DEZEMBRO/2019

GLEYSONN MENANDRO DE SOUZA

**SPORTS ANALYTICS: VARIÁVEIS SIGNIFICANTES NA SELEÇÃO DE
QUARTERBACKS NA NFL**

**Monografia apresentada ao Curso de Graduação
em Gestão da Informação, da Faculdade de
Gestão e Negócios da Universidade Federal de
Uberlândia, como exigência parcial para a
obtenção do título de Bacharel.**

Orientador: Prof. Dr. Élcio Eduardo de Paula Santana

**UBERLÂNDIA-MG
DEZEMBRO/2019**

SPORTS ANALYTICS: VARIÁVEIS SIGNIFICANTES NA SELEÇÃO DE QUARTERBACKS NA NFL

Monografia aprovada para a obtenção do título de Bacharel no Curso de Graduação em Gestão da Informação da Universidade Federal de Uberlândia pela banca examinadora formada por:

Uberlândia, 09 de Dezembro de 2019.

Prof. Dr. Élcio Eduardo de Paula Santana, FAGEN/UFU

Prof. Dr. André Francisco Alcântara Fagundes, FAGEN/UFU

Prof. Dr. José Eduardo Ferreira Lopes, FAGEN/UFU

RESUMO

Dado a importância do *quarterback* para os times da liga de futebol americano - a NFL – dos Estados Unidos da América, é cada vez mais importante saber quais as variáveis devem-se levar em consideração na hora do *draft*. Esta pesquisa visa verificar quais variáveis, dentre as mensuradas ao longo da carreira na universidade e no combine, predizem o *rating* do *quarterback* na NFL. A pesquisa foi realizada com dados oficiais coletados de fontes secundárias e utilizou-se da análise de dados multivariados, especificamente a regressão múltipla, para medir a relação entre a variável dependente e as independentes. Verificou-se que as variáveis de desempenho levadas em consideração atualmente na seleção de *quarterbacks* não apresentam significância na previsão do desempenho dos atletas no profissional. A partir dos resultados apresentados conclui-se que o desempenho do *quarterback* no *Combine* e seu histórico de performance nos jogos do *College* não influenciam de forma significativa seu desempenho na NFL. Evidenciando que analisar apenas variáveis de desempenho ao recrutar um *quarterback* pode-se não ser suficiente.

Palavras-chave: *Sports Analytics; NFL; Draft; Combine; Quarterbacks; Estatística; Marketing Esportivo; Atleta como produto*

ABSTRACT

Given the importance of the quarterback for the football league teams - the NFL - from the United States of America, it is increasingly important to know which variables should be taken into account at the time of the draft. This research aims to verify which variables, among those measured throughout his career at the university and in the combine, predict the quarterback's rating in the NFL. The research was conducted with official data collected from secondary sources and multivariate data analysis was used, specifically multiple regression, to measure the relationship between the dependent variable and the independent variables. It was found that the performance variables currently taken into account in the selection of quarterbacks do not present significance in predicting the performance of athletes in the professional. From the results presented it is concluded that the quarterback's performance at the Combine and his performance history in College games do not significantly influence his performance in the NFL. Evidencing that analyzing only performance variables when recruiting a quarterback may not be enough.

Keywords: *Sports Analytics; NFL; Draft; Match; Quarterbacks; Statistic; Sports marketing; Athlete as a product*

LISTA DE FIGURAS

Figura 1 – Ranking das Ligas Esportivas por Receita

Figura 2 – Metodologia do *40-yard Dash*

Figura 3 – Metodologia do *20-yard Shuttle*

Figura 4 – Metodologia do *3-Cone Drill*

Figura 5 – Metodologia do *Vertical Jump*

Figura 6 – *Broad Jump*

Figura 7 – *Bench Press Combine*

Figura 8 – *Sports Analytics Framework*

Figura 9 – Seleção de uma técnica Multivariada

Figura 10 – Análise gráfica de resíduos

LISTA DE GRÁFICOS

Gráfico 1 – Dispersão dos Dados

Gráfico 2 – Distribuição normal

Gráfico 3 – Distribuição normal dos Resíduos padronizados

LISTA DE TABELAS

Tabela 1 – Variáveis Dependentes e Independentes

Tabela 2 – Estatística Descritiva

Tabela 3 – Correlações

Tabela 4 – Estatísticas de Colinearidade

Tabela 5 – Estatísticas de Resíduos

Tabela 6 – Resumo do Modelo

Tabela 7 – Teste de significância

Tabela 8 - Coeficientes

LISTA DE ABREVIATURAS E SIGLAS

AFC – American Football Conference

AFL - American Football League

APFA – American Professional Football Association

Att - Attempt

BCS - Bowl Championship Series

CEO - Chief Executive Officer

Comp - Complete

ESPN - Entertainment and Sports Programming Network

FBS - Football Bowl Subdivision

FCS - Football Championship Subdivision

G - Games

In - Inches

INT - interception

Lbs - Pounds

MLB – Major League Baseball

NASA - National Aeronautics and Space Administration

NBA – National Basketball League

NCAA - National Collegiate Athletic Association

NFC – National football Conference

NFL – National Football League

NIC - National Invitational Camp

QB - Quarterback

QI – Quociente de Inteligência

RAT - Rating

TD - Touchdowns

VIF – Variance Inflation Factor

Yds - yards

LISTA DE APÊNDICES

Apêndice A – Modelo 1: Rating com 16 variáveis independentes com multicolinearidade

Apêndice B – Modelo 2: Rating com 12 variáveis independentes

Apêndice C – Modelo 3: Draft

SUMÁRIO

1. INTRODUÇÃO	1
2. REFERENCIAL TEÓRICO	4
2.1. Atleta como Produto	4
2.2. A Importância do Recrutamento	5
2.3. NFL	8
2.4. Draft	10
2.5. Combine	11
2.5.1. 40-Yard Dash	12
2.5.2. 20-Yard Shuttle	12
2.5.3. 3-Cone Drill	13
2.5.4. Vertical Jump	14
2.5.5. Broad Jump	14
2.5.6. 225-Bench Press	15
2.5.7. Wonderlic	16
2.6. Quarterback	16
2.7. Sports Analytics	17
3. ASPECTOS METODOLÓGICOS	21
3.1. Definição das Variáveis	21
3.2. Técnica e Instrumento de Coleta	22
3.3. Definição da Técnica de Análise de Dados	23
4. DESCRIÇÃO E ANÁLISE DOS DADOS	26
4.1. Análises Preliminares	26
4.2. Poder Estatístico e Tamanho da amostra	26
4.3. Pressupostos da Análise de Regressão Múltipla	29
4.3.1. Multicolinearidade	29
4.3.2. Outliers	33
4.3.3. Normalidade	34
4.3.4. Linearidade	35
4.3.5. Heterocedasticidade	35
4.3.6. Independência	36
4.3.7. Estimação do Modelo e Avaliação do Ajuste Geral	37
4.3.7.1. Seleção da Técnica de Estimação	37

4.3.7.2.	Avaliação do Modelo	38
4.3.8.	Avaliando as Variáveis Independentes	39
5.	CONSIDERAÇÕES FINAIS	41
5.1.	Contribuições da Pesquisa	41
5.2.	Sugestões para Pesquisas Futuras	42
5.3.	Limites do Estudo	42
	APÊNDICE A – Rating com 16 variáveis independentes com multicolinearidade.....	47
	APÊNDICE B – Rating com 12 variáveis independentes	57
	APÊNDICE C – Draft	66

1. INTRODUÇÃO

O futebol americano teve seu início com o surgimento das associações e com rugby, de origem britânica, embora o esporte não tenha muita relação com o que conhecemos hoje, seu primeiro jogo aconteceu na faculdade em 1869 entre Rutgers e Princeton (HENNINGER, 2018).

Os clubes atléticos se tornavam cada vez mais populares depois da Guerra civil americana, assim as rivalidades começaram a surgir, e em 12 de novembro de 1892 aconteceu um jogo que mudou o futebol americano. Nesta data a *Allegheny Athletic Association* derrotou o *Pittsburg Athletic Club*, onde William Heffelfinger recebeu US \$ 500 para participar do jogo, se tornando o primeiro atleta a ser pago para jogar futebol americano da história (HENNINGER, 2018).

Durante a 1ª Guerra Mundial os clubes começaram a recrutar jogadores de estados diferentes, devido a dificuldade em se encontrar homens atléticos para jogar localmente, incidindo na necessidade de uma regulamentação mais eficaz, levando a criação da *American Professional Football Association* (APFA), em 1920 com 14 equipes (HENNINGER, 2018).

Dois anos após a formação da APFA, a liga alterou seu nome para *National Football League* (NFL). O ponto de ruptura para o esporte moderno aconteceu após os playoffs de 1932, a incorporação da AP Poll¹, a criação do *Heisman Trophy Award*², o primeiro *draft* e o primeiro jogo televisionado. Essas mudanças transformaram a NFL no que conhecemos atualmente (HENNINGER, 2018).

O esporte começou então a atrair fãs e espectadores por todo o território americano. A popularidade fez com que surgisse, em 1929, uma nova liga, a *American Football League* (AFL). Em 1960 a liga já contava com 8 equipes e a pressão para que acontecesse uma fusão entre a AFL e a NFL aumentou (HENNINGER, 2018). A fusão veio a acontecer em 1970, surgindo novas equipes e mudanças nas regras (HENNINGER, 2018).

¹ Pesquisa da Associated Press (AP Poll) fornece classificações semanais das 25 principais equipes da NCAA, a pesquisa pode ser acessada em <https://apnews.com/APTop25CollegeFootballPoll>.

² O Troféu Heisman, nomeado em homenagem ao ex-jogador da Brown University e treinador John Heisman, é um prêmio anual dado pela Heisman Trophy Trust para o melhor jogador da temporada do futebol americano universitário, a lista completa de ganhadores pode ser acessada no site: <https://www.heisman.com/heisman-winners/>.

O esporte então veio a se tornar muito lucrativo, e o sonho de parte dos atletas de ensino médio, em 2016 a liga faturou estimados US\$ 13,3 bilhões de dólares, em 2017 esse número sobe para 14,2 bilhões e a Bloomberg estima que em 2018 a NFL faturou cerca de 15 bilhões de dólares. O comissário Roger Goodell alvejou US \$ 25 bilhões em receita até 2027, ou seja, 6% de crescimento anual (ECKSTEIN, 2019).

A NFL agrupa seus fluxos de receita em duas categorias: receita nacional e receita local (ECKSTEIN, 2019). A receita nacional consiste de acordos de TV, merchandising e licenciamentos negociados pela própria NFL. O dinheiro é dividido entre as 32 equipes, independentemente do desempenho individual. De acordo com o relatório anual do Green Bay Packers – único time da NFL de capital aberto – a liga faturou mais de US\$ 8,1 bilhões em receita nacional em 2018, cada equipe recebendo a quantia de US\$ 255 milhões (ECKSTEIN, 2019).

A receita local consiste da venda de ingressos, concessões e patrocinadores corporativos, e é obtida pelas próprias equipes. Em 2018, os Packers ganharam US \$ 196 milhões em receita local, 43% de sua receita total naquele ano, que foi de US \$ 455 milhões (ECKSTEIN, 2019).

No entanto, Eckstein (2019) ainda menciona que os custos da administração de um time de futebol americano são altos. Em 2018, os Packers gastaram 420 milhões de dólares em despesas, US\$ 213 milhões para jogadores e US\$ 208 milhões para manutenção de estádios, marketing e custos administrativos, deixando um lucro operacional de US\$ 38,5 milhões. O Dallas Cowboys é a equipe mais rica da NFL, com US\$ 864 milhões em receita e lucro operacional de US\$ 365 milhões em 2018 (ECKSTEIN, 2019). Em 2019, segundo a Forbes (2019), esse valor já chega a US\$ 950 milhões em receita e US\$ 420 milhões em lucro operacional.

Entre os atletas mais bem pagos em 2019, de todos os esportes, o *quarterback (QB)* Russel Wilson do Seattle Seahawks está em sexto lugar, com salário anual de US\$ 89,5 milhões, seguido da estrela do Green Bay Packers, Aaron Rodgers, com faturamento anual de US\$ 89,3 milhões de dólares, colocando os atletas da NFL entre os esportistas mais bem pagos do mundo (FORBES, 2019). Tom Brady, *QB* com maior número de Superbowls, está somente na 81ª colocação com salário de US\$ 27 milhões de dólares.

Identificando assim que a NFL possui o maior faturamento entre as ligas (Figura 1, pag. 21) e, dentre todos os atletas que participam de um time, o *QB* desempenha um papel primordial para o jogo, sendo o atleta em campo que deve conhecer todo o *playbook* de jogadas e o responsável por inicia-las de acordo com a coordenação ofensiva. De suma

importância para o marketing das organizações e ainda, dentre todos os esportistas, aqueles entre os 10 mais bem pagos, este trabalho é de suma importância tanto academicamente, pois visa ampliar os estudos sobre o tema, e quanto gerencialmente, pela verificação de quais variáveis predizem o *rating* do *quarterback* na NFL.

Desta forma dados sobre os atletas serão analisados por meio de modelos estatísticos a fim de verificar quais variáveis, dentre as mensuradas ao longo da carreira na universidade e no *combine* do *draft*, vaticinam o *rating* do *quarterback* na NFL.

2. REFERENCIAL TEÓRICO

Neste item será abordado a importância do recrutamento e dado ênfase no âmbito esportivo. Será analisado as formas que os principais esportes americanos lidam com o recrutamento e seleção de atletas. A história da NFL (*National Football League*), a importância do *quarterback* para as equipes e as ferramentas de seleção de jogadores utilizada na liga americana de futebol americano.

2.1. ATLETA COMO PRODUTO

A História provou que os atletas podem vender produtos, e talvez mais importante, criar conexões entre marca e consumidor (ARTHUR, 2016), se tornando uma estratégia essencial para as grandes marcas do esporte.

Em Wimbledon quando Djokovic perdeu um desempate para Federer, tentou e falhou em arrancar sua própria camiseta gerando um marketing único sobre a qualidade das roupas fabricadas pela Uniqlo – marca japonesa de roupas (ARTHUR, 2016).

Em 1985, a Nike contratou o então novato Michael Jordan como porta-voz. Jordan ainda era uma promessa, mas ele personificava um desempenho superior (KOTLER; KELLER, 2013). A aposta da Nike compensou: a linha Air Jordan de tênis para basquete sumiu das prateleiras e as receitas atingiram mais de US\$ 100 milhões somente no primeiro ano (KOTLER; KELLER, 2013).

A história se repete para Tommy Hilfiger, que contratou Rafael Nadal como embaixador global da marca de roupas íntimas, alfaiataria e fragrância. O resultado segundo o chefe de marca e diretor de marketing, as vendas de roupas íntimas dobraram naquele ano, atribuindo o sucesso ao envolvimento de Nadal (ARTHUR, 2016).

Muitas marcas têm como alvo os esportes de mercado de massa para acessar um conjunto mais amplo de consumidores - o futebol, com seus estimados 3,5 bilhões de fãs globais, é um deles (ARTHUR, 2016).

Um exemplo foi o efeito Neymar. O PSG (Paris-Sant Germain) ao contratar o brasileiro Neymar teve que arcar com uma multa rescisória do Barcelona de 221 milhões de euros, porém a venda de camisetas com o nome do atleta vendeu 1 milhão de euros em um dia, sendo que parte considerável do valor fica com o fabricante, ou seja, a Nike (SALVADOR, 2017).

Outras marcas também demonstram essa estratégia. Belstaff trabalha com David Beckham, Ralph Lauren negocia com o jogador de pólo Nacho Figueras, e a empresa de malas Tumi assinou no ano passado o piloto de Fórmula 1 Nico Rosberg. As marcas de relógios também dominam: Nadal usa Richard Mille; Federer e Tiger Woods trabalham com a Rolex; Lionel Messi e Serena Williams com Audemars Piguet; Rory McIlroy com Omega; Usain Bolt, Jérôme Boateng do Bayern de Munique e Pelé com a Hublot; e Cristiano Ronaldo e o *quarterback* seis vezes campeão do *super bowl*, Tom Brady com a TAG Heuer (ARTHUR, 2016).

Os atletas sempre foram uma das maneiras mais diretas de alcançar os consumidores. O esporte desperta paixões, é fonte de saúde, sacrifício, comprometimento, positivismo e, levado ao nível de elite, desenvolve o fenômeno dos fãs que tem poucos concorrentes em outros setores da sociedade (CRUYFF, 2015).

As redes sociais também tem mudado os conceitos de marketing, não sendo mais necessário que a empresa necessite fazer uma megaprodução para conseguir atingir o público alvo desejado. O atleta entrega ao patrocinador milhões de pessoas que estão interessadas no que ele faz, economizando tempo (CRUYFF, 2015).

O astro do futebol 5 vezes eleito como melhor do mundo, Cristiano Ronaldo obteve rendimento com redes sociais maior que seu contrato com a Juventus (LANE, 2019). Com rendimento de 47,8 milhões de dólares em postagens pagas do Instagram em 2018, 14,7 milhões a menos. Entre as marcas que Ronaldo representou no Instagram estão a Clear Haircare, Nike Football e Six Pad Europe (LANE, 2019).

No NFLshop.com, a estrela do New England Patriots, e marido da brasileira Gisele Bündchen, Tom Brady lidera a venda de *jerseys* em 2019, seguido de Patrick Mahomes *quarterback* do Kansas City Chiefs - décima escolha da primeira rodada do *draft* de 2017.

2.2. A IMPORTÂNCIA DO RECRUTAMENTO

Sabe-se que no mundo organizacional o recrutamento e seleção de funcionário se tornou uma forte ferramenta de gestão capaz de garantir as competências necessárias à organização. Segundo Carvalho (2008), o objetivo de recrutamento e seleção de profissionais deixou de ser para o simples fato de se preencher cargos, estando voltados mais para a contratação de profissionais sintonizados com as necessidades estratégicas da empresa.

Com a necessidade de evoluir os processos de recrutamento e seleção, veio o desenvolvimento de uma ciência de estudo sobre o comportamento dos acontecimentos organizacionais, a psicologia organizacional. Uma das contribuições mais expressivas da psicologia organizacional aos processos de seleção foi a utilização de testes como instrumentos de medição dos traços psicológicos julgados necessários para desempenhar funções nos respectivos cargos (CARVALHO, 2008).

Com as mudanças e necessidades empresariais modernas os métodos de recrutamento e seleção ganharam um papel importante no meio organizacional. Esses métodos - como dinâmicas em grupos, provas e entrevistas - visam deixar o processo de seleção de profissionais mais vantajoso para a organização de forma a possibilitar uma maior capacidade de análise do desempenho dos candidatos. No mundo esportivo não é diferente, no caso do futebol, que segundo Marcos Guterman (2010) é o maior fenômeno social do Brasil, representando uma identidade nacional e dando objetivo aos desejos dos brasileiros, é observado a habilidade técnica e tática do atleta, assim como suas condições físicas e psicológicas sustentados por cinco elementos no processo de detecção: I) Habilidade técnica; II) Inteligência de jogo; III) Equilíbrio Emocional; IV) Condição física; e V) Regras do jogo (PAOLI, SILVA e SOARES, 2008, p. 43).

Assim como nas empresas, nos esportes as técnicas de recrutamento visam a identificação dos melhores perfis para cada posição, a fim de se obter atletas com habilidades que venham a ser desenvolvidas ou que estejam em nível profissional:

A detecção de talentos no futebol tem como fim identificar atletas com potencial para o alto nível de desempenho. Tal processo é iniciado nas categorias de base, de maneira a contribuir para formá-los e torná-los atletas competitivos, capazes de obter bons resultados em competições, e que possam atingir a categoria profissional e, se tornar um bom “produto” para negócios. (PAOLI, SILVA E SOARES, 2008, p. 43)

As seleções de atletas nos esportes americanos buscam promover o equilíbrio entre as equipes, desta forma na NBA (*National Basketball Association*) o processo de recrutamento de jogadores é realizado por um *draft* – um evento anual no qual 30 times podem recrutar jogadores que são elegíveis para ingressar na liga, sejam universitários ou estrangeiros (ANDRADE, 2019).

São selecionados 60 atletas divididos em duas rodadas, e a definição da ordem de escolha é feita por um sorteio, conhecido como *Draft Lottery*, não existindo garantia de que o pior time da temporada anterior da NBA terá a primeira escolha do draft, o que acontece é que a equipe tem uma maior chance de recrutar primeiro (ANDRADE, 2019).

No *Draft Lottery* participam 14 times que não se classificaram para os *playoffs*, os outros 16 times ficam com as escolhas seguintes, seguindo uma ordem de acordo com as campanhas na temporada regular (os *playoffs* são considerados) (ANDRADE, 2019). O sorteio apenas define a ordem das 3 primeiras escolhas, as demais seguem a ordem da classificação na temporada (ANDRADE, 2019).

Porem as equipes ainda podem negociar as suas escolhas no *draft*, e ainda há a possibilidade de os jogadores serem trocados em seguida (ANDRADE, 2019). Um exemplo foi o brasileiro Nenê, draftado pelo New York Knick na sétima escolha, mas negociado com o Denver Nuggets em seguida (ANDRADE, 2019). Michael Jordan, considerado o melhor jogador de todos os tempos, foi apenas a terceira escolha do *draft* de 1995 (ANDRADE, 2019).

O ano de 2018 foi o último da loteria no sistema de recrutamento da NBA, a partir de 2019 os três times com as piores campanhas terão as mesmas chances de levar a 1ª escolha (14%) (ANDRADE, 2019). A iniciativa nasceu para tentar combater o *tanking* – quando o time tenta perder mais jogos para ter maiores chances de ficar com as primeiras escolhas (ANDRADE, 2019).

O *draft* na MLB (*Major League Baseball*) é o mais longo de todos os esportes americanos. Cada time seleciona em cada uma das rodadas, tornando extenso o número de jogadores draftados (VEIGA, 2017). Ao todo são escolhidos mais de mil atletas todos os anos (VEIGA, 2017).

Ainda que a quantidade de jogadores draftados seja grande não são todos que assinam contrato com as equipes, o atleta tem um prazo, normalmente, de um mês para assinar com a equipe que o escolheu (VEIGA, 2017). A equipe apresenta um "Bônus" (em dinheiro) e cabe ao atleta aceitar ou não, caso ele não aceite ele ficará inelegível para assinar com qualquer time naquele draft em questão, podendo tentar uma nova chance no ano seguinte (VEIGA, 2017).

Sendo muito comum entre atletas no último ano do ensino médio, onde arriscam serem draftados na expectativa de receber um bônus grande, ou cair em um time que julgam ser o certo (VEIGA, 2017). Caso as condições não sejam de seu interesse os atletas optam por recusar a oferta e entrar em uma universidade (VEIGA, 2017).

Entre as equipes existe a possibilidade de deixar escolhas passarem, não sendo obrigatório que selecionem algum jogador, podem simplesmente desistir daquela *pick* em questão (VEIGA, 2017). Com relação a ordem de escolha, não se difere da NFL, porém

não há loteria como na NBA, o primeiro a escolher é o time de pior campanha, facilitando o *tanking* (VEIGA, 2017).

2.3. NFL

O futebol americano migra para os Estados Unidos por meio de jovens americanos que iam estudar em Londres e lá tinham contato com o *Association Football* e o *Rugby Football* se espalhando pelas escolas e clubes americanos na década de 60 (ESPN, 2016) e sendo jogado pela primeira vez em 1869, por Rutgers e Princeton, com regras modificadas do *Association Football* e *Rugby Football*. Walter Camp, na convenção de Massasoit (NFL, 2016), escreveu as primeiras regras do American Football se tornando “o pai do Futebol Americano”, inserindo as regras dos *downs* – quando um jogador é derrubado o árbitro paralisa o jogo para que os times se reagrupem possuindo chances para que andem determinada quantidade de jardas (ESPN, 2016).

Em 1905 o futebol americano quase foi proibido pelo então presidente Roosevelt devido ao grande número de lesões, cominando na revisão das regras do esporte, a fim de torna-lo mais seguro. É nesse período que o *Forward Pass* (Passe para frente), foi incluído nas regras, ganhando a popularidade nos Estados Unidos (ESPN, 2016).

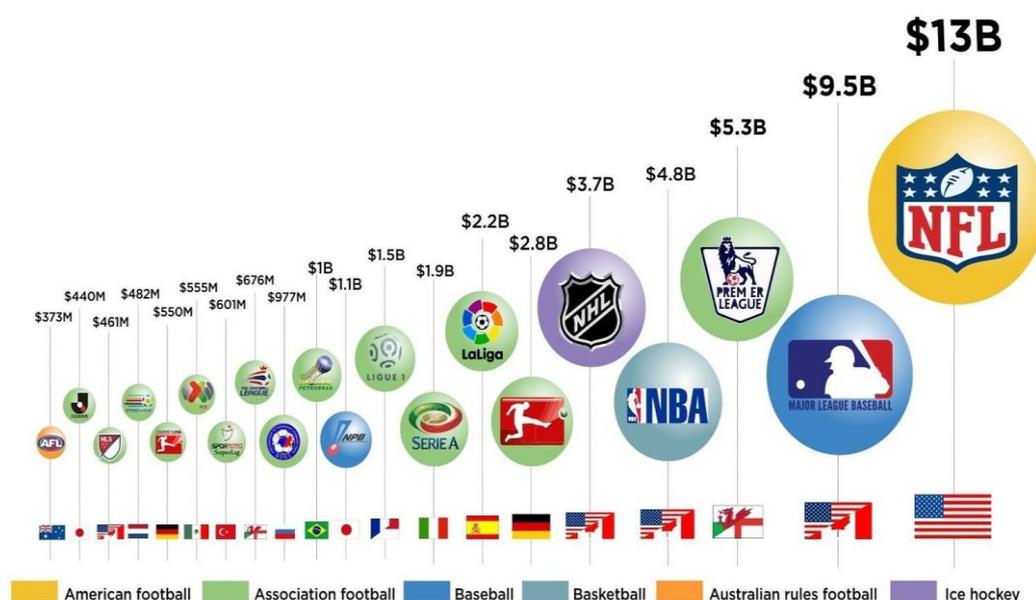
De acordo com a NFL (2016), existia problemas no campeonato profissional: os jogadores começaram a receber uma enorme quantia em salários, ficavam saindo das equipes por receberem propostas melhores e os times usavam jogadores universitários – caracterizados como aqueles que ainda estão matriculados em uma universidade, assim, no intuito de realizar um campeonato mais igualitário, do qual todos seguissem as mesmas regras, foi realizada um conferencia em Ohio com profissionais que resultou na *American Professional Football Conference* (Conferencia Profissional de Futebol Americano) sendo alterado posteriormente para *American Professional Football Association* (Associação Profissional de Futebol Americano) a atual NFL (*National Football League*), alterando seu nome em 1922.

A NFL começou a recusar a entrada de outros times na liga, o que deu origem a AFL (*American Football League*), que inovou o esporte criando a conversão de dois pontos – chance que um time pode ter de realizar um passe ao invés de chute de ponto, tentando um novo touchdown – e o marketing com nomes de atletas nas camisas (ESPN, 2016). A AFL foi a única liga a conseguir concorrer com a NFL, e após muitos anos de tensão entre as duas concorrentes, foi criado o *Super Bowl* em 1967, uma final entre o

campeão da AFL e da NFL para definir o campeão nacional de futebol americano, do qual naquele ano o campeão foi o Green Bay Packers. Atualmente o *Super Bowl* é um evento mundialmente conhecido, gerando uma receita no *Super Bowl 50* entre Denver Broncos e Carolina Panthers de aproximadamente US\$ 620 milhões segundo a Forbes (OZANIAN, 2016).

Hoje a NFL conta com 32 times, divididos em duas ligas AFC e NFC sendo projetado para o ano, segundo Roger Goodell (comissário da NFL), ultrapassar uma receita total de US\$13.3 Bilhões (BELZER, 2016), a maior receita de uma liga esportiva do mundo.

Figura 1 – Ranking das Ligas Esportivas por Receita



Fonte: Retirado do site *How Much*³

No ano de 1936 teve início o draft universitário – um recrutamento do qual os times são colocados em ordem inversa, ou seja, o pior time da liga é o primeiro a escolher um jogador – mantendo sempre a competitividade na liga (NFL, 2016).

³ Disponível em: < <http://howmuch.net/articles/sports-leagues-by-revenue> >. Acessado em: 08 de Julho de 2016

2.4. DRAFT

O *Draft* ocorre durante três dias a cada primavera. O *Draft* da NFL dá aos times a opção de trazer novos atletas possuindo 10 minutos para fazer sua escolha na primeira rodada, sete minutos das rodadas 3 a 6 e apenas 4 minutos na última rodada e em *picks* extras, caso a equipe não realize sua escolha no durante seu tempo ela poderá realizar posteriormente, mas dá a oportunidade da equipe subsequente de selecionar um jogador que seria sua escolha naquela rodada (NFL OPERATION, 2016).

As equipes recebem cada uma 7 (sete) escolhas, uma em cada rodada do NFL *Draft*, onde o primeiro clube a escolher é o de pior campanha da temporada anterior e o último é o campeão do *Super Bowl*. Caso ocorra empate de desempenho na temporada o critério de seleção será a porcentagem de vitórias sendo aquela que tiver o pior resultado terá o melhor lugar perante a(s) outra(s) equipe(s), e se ainda assim houver empate será levado em consideração seus registros contra adversários comuns. Se não for possível aplicar esses critérios ou se o empate persistir em todos será jogado na moeda (NFL OPERATION, 2016).

A NFL permite a comercialização de seus prospectos escolhidos e qualquer tipo de negociação de jogadores que a equipe já possui direitos e ainda mantém um sistema de *picks* extras ou compensatórios para aquelas equipes que perderam jogadores para outros clubes. A organização calcula, por meio de uma fórmula patenteada, o valor que a equipe perdeu baseada nos salários dos jogadores que saíram e o tempo jogado, sendo dado a equipe no máximo 4 *picks* entre o final da terceira rodada e sétima rodada, totalizando o valor líquido perdido (NFL OPERATION, 2016).

Para se eleger para o *Draft* o atleta deve ter concluído o ensino médio por pelo menos três anos e deve ter utilizado sua elegibilidade antes da próxima temporada de futebol americano universitário (NFL OPERATION, 2016), que é a referência de novos atletas a participarem do *Combine* e do *Draft*. Os atletas que terminaram a faculdade e não utilizaram sua elegibilidade podem realizara a solicitação de entrar no *Draft* junto a Liga, sendo elegíveis apenas no ano após o fim de sua elegibilidade (NFL OPERATION, 2016).

A NFL trabalha junto a NCAA – *National Collegiate Athletic Association* – a fim de verificar a elegibilidade dos prospectos e autorizar a participação no *Draft* (NFL OPERATION, 2016). A NCAA é uma associação esportiva das faculdades americanas dedicadas aos atletas universitários, onde mais de 460 mil atletas universitários compõem

as 19 mil equipes que enviam mais de 54 mil participantes para competir a cada ano em 90 campeonatos da NCAA em 24 esportes diferentes por 3 divisões (NCAA, 2016).

No futebol americano a NCAA é dividida em quatro níveis: *Football Bowl Subdivision* (FBS), *Football Championship Subdivision* (FCS), *Division II* e *Division III*. A FBS é a principal divisão de futebol americano da NCAA, com 124 equipes, divididas em 11 conferencias (LILLY, 2012).

O objetivo dos times universitários é competir nas temporadas regulares e serem campeões ou vice-campeões de suas respectivas conferencias, os campeões das 6 ligas principais mais 4 times são convidados para o *BCS Bowl Games*, que inclui o *Orange Bowl* em Miami, o *Fiesta Bowl* em Phoenix, o *Sugar Bowl* em New Orleans, o *Rose Bowl* em Pasadena e o *BCS National Championship Game* do qual o campeão é nomeado *BCS Champion* ou o campeão nacional universitário (LILLY, 2012).

2.5. COMBINE

O *Combine* é um grande evento de testes médicos, psicológicos e físicos - além de entrevistas com executivos, treinadores e olheiros - que tem como objetivo determinar se o atleta irá para o próximo *Draft*. Conta com a participação de mais de 300 prospectos a jogadores anualmente. O *NFL Football Scouting* foi organizado pela primeira vez em 1982 quando o departamento da NFL decidiu criar o *National Invitational Camp* - NIC, realizando os testes com apenas 163 prospectos em Tampa, na Flórida, hoje realizado em Indianapolis, Indiana.

O evento NIC ganhou força com o tempo quando os times da NFL concordaram em participar do *Combine*, sendo rebatizado de *NFL Scouting Combine* em 1985. No mundo do futebol americano tem-se o que melhor existem em treinadores, o papel então dos olheiros é encontrar os melhores jogadores possíveis para sua estrutura e seus objetivos, assim gastam muito de seu tempo vendo vídeos de prospectos jogadores. O *Combine* permite o agrupamento, testes de alto nível e monitoramento desses prospectos em um lugar, durante 4-5 dias (ROBERTSON, 2016).

São seis testes que todos os prospectos são submetidos no *Scouting Combine*: *40-yard dash*, *20-yard shuttle*, *3-cone drill*, *vertical jump*, *broad jump*, e *225-Bench Press*.

2.5.1. 40-YARD DASH

No tiro de 40 jardas⁴, o atleta deve se posicionar agachado e permanecer por pelo menos 2 segundos antes de começar a correr. O atleta deve correr em linha reta durante 40 jardas sendo cronometrado por mecanismos ao lado da pista ao atingir 10, 20 e 40 jardas (ROBERTSON, 2016).

Este teste pode ser dividido em três partes para determinar o perfil do atleta: 0-10 jardas, é considerada a zona de aceleração; 11-25 jardas, é a zona de máxima velocidade; e de 26-40 jardas, é a zona de resistência. Cada zona fornece indicadores de força, saltos, explosão e velocidade máxima, sendo bons indicadores tanto para atletas mais rápidos quanto para atletas mais lentos, uma vez que pode ser utilizado para medir a resistência de atletas mais pesados (ROBERTSON, 2016).

Figura 2 – Metodologia do 40-yard Dash



Fonte: Modificado de *Hardcore training*⁵

2.5.2. 20-YARD SHUTTLE

No *20-yard Shuttle* o atleta se posiciona agachado na posição central de 10 (dez) jardas, começa correndo para a direita durante 5 (cinco) jardas e obrigatoriamente deve tocar a linha com a mão, correndo de volta durante as dez jardas tocando com a mão novamente a linha e voltando para a posição inicial totalizando 20 jardas, é um indicador de capacidade que o atleta possui de alterar sua direção em uma corrida (ROBERTSON, 2016).

⁴ Uma jarda equivale a 91,44 centímetros.

⁵ Disponível em: < <http://www.hardcore40training.net/measure.php#> >. Acessado em: 30 de novembro de 2019.

Figura 3 – Metodologia do 20-yard Shuttle



Fonte: *Hardcore Training*⁶

2.5.3. 3-CONE DRILL

São colocados três cones formando um “L”, o prospecto começa em posição agachada do cone da extremidade do “L” e corre para o cone à sua frente, volta para o cone de início e toca a linha de partida com a mão, volta ao cone central contornando a 90 graus correndo ao terceiro cone, contorna o cone e retorna ao início (ROBERTSON, 2016), conforme imagem abaixo:

Figura 4 – Metodologia do 3-Cone Drill



Fonte: *Hardcore Training*⁷

Este teste é o que mais se parece com os movimentos encontrados em um jogo real de futebol americano, é um indicador de aceleração, desaceleração e mudança de

⁶ Disponível em: < <http://www.hardcore40training.net/measure.php#> >. Acessado em: 30 de novembro de 2019.

⁷ Disponível em: < <http://www.hardcore40training.net/measure.php#> >. Acessado em: 30 de novembro de 2019.

direção, onde o atleta deve se movimentar rapidamente e realizar bruscas mudanças de direção (ROBERTSON, 2016).

2.5.4. VERTICAL JUMP

Salto vertical é um indicador de explosão da parte inferior do corpo. O atleta fica em pé e estica o braço obtendo a primeira medição (M1), então o atleta salta tocando na bandeira obtendo então a segunda medição (M2), a diferença entre a segunda medição e a primeira é o Salto Vertical do atleta (NFL WORKOUTS, 2016).

$$\textit{Vertical Jump} = \textit{Medida2} - \textit{Medida1}$$

Figura 5 – Metodologia do *Vertical Jump*



Fonte: *Flight Basketball*⁸

2.5.5. BROAD JUMP

No *Broad Jump* o atleta se posiciona e salta para frente, não podendo se mexer ao pousar, onde é medido seu desempenho no teste, este salto é um indicador de força da parte inferior do corpo e de equilíbrio (NFL WORKOUTS, 2016).

⁸ Disponível em: <<http://www.flightbasketball.com/best-vertical-jump-workout-program-explosive-results/>>. Acessado em: 09 de Julho de 2016.

Figura 6 – *Broad Jump*

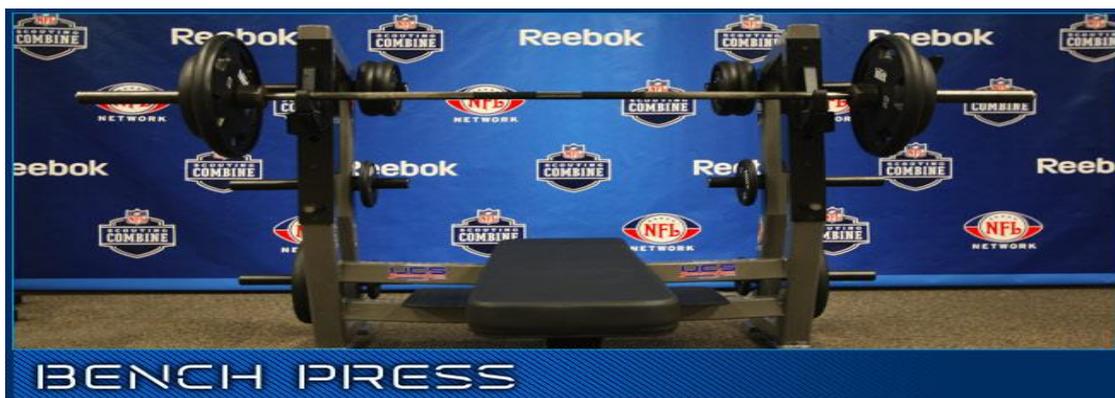


Fonte: *Sporting News*⁹

2.5.6. 225-BENCH PRESS

O 225 – *Bench Press* ou simplesmente supino é um teste de repetições com 225 libras de peso do qual procura-se medir a resistência do prospecto atleta (NFL WORKOUTS, 2016). Cada atleta pode fazer o teste apenas uma vez, a barra deve tocar o peito, os cotovelos devem ficar eretos no topo, o corpo não pode levantar ao realizar o movimento e o teste é realizado até a falha muscular (NFL UP, 2016).

Figura 7 – *Bench Press Combine*



Fonte: *Creativity Post*¹⁰

⁹ Disponível em: < <http://www.sportingnews.com/nfl/news/nfl-combine-records-history-40-time-bench-press-cone-drills/87nkjnk222my1y1oip119lqu1>>. Acessado em: 09 de Julho de 2016.

¹⁰ Disponível em: < http://www.creativitypost.com/psychology/how_do_you_create_an_elite_athlete>. Acessado em: 09 de Julho de 2016.

2.5.7. WONDERLIC

O teste Wonderlic foi desenvolvido por E.F. Wonderlic em 1936, assim como um teste de QI (quociente de inteligência) o Wonderlic foi desenvolvido para mensurar a inteligência dos prospectos no Combine. Sendo uma prova de 12 minutos com 50 questões, um resultado médio equivale a 20 pontos sendo o mesmo que um QI de 100, enquanto os prospectos podem ter uma pontuação máxima de 50 (WONDERLIC TEST SAMPLE, 2019)

O teste não foi usado diretamente na NFL, uma de suas primeiras aplicações foi na marinha americana, onde foi utilizado para testar a inteligência dos pilotos colocando-os sob pressão. Foi usado pela primeira vez na NFL na década de 70 pelo então técnico do Dallas Cowboys, Tom Landry (WONDERLIC TEST SAMPLE, 2019).

Segundo o NFL *Combine Results* o *quarterback* que obteve a pontuação mais baixa foi de Oscar Davenport que tirou 6 em 1999. Ryan Fitzpatrick tirou incríveis 48 no teste em 2005 sendo essa a pontuação mais alta e explicando o porquê de estudar em Harvard, enquanto que Tom Brady tirou 33 no wonderlic e ganhou seis *super bowls*.

2.6. QUARTERBACK

O *Quarterback* (QB) é sem dúvida a posição mais importante do futebol americano, responsável por todas as jogadas e por comandar o time dentro de campo, sendo exigido e cobrado durante toda a temporada (FORLIN, 2013). É o cérebro do time, realiza as jogadas lançando a bola em profundidade, correndo com a bola ou entregando-a a outro jogador em um passe curto (ESPN, 2016), se posiciona atrás da linha de ataque que serve como proteção, segurando a linha defensiva da outra equipe até o QB realizar a jogada. É a peça chave do ataque de um time de futebol americano (ESPN, 2016) determinando por meio de seu controle da partida um jogo e até mesmo um campeonato.

Dado sua importância, no *Draft* de 2014 o Cleveland Browns tinha dois *picks* de primeira rodada e precisava de um QB. Os Browns encomendaram uma análise de US\$ 100.000, o estudo analisava quais eram as qualidades necessárias para um QB ter sucesso na liga (BONESTEEL, 2016). Após análise dos resultados, o estudo previu que Teddy Bridgewater era melhor *quarterback* que Johnny Manziel, porém os Browns ignoraram esta informação e selecionaram Manziel no 22º *pick* da primeira rodada e Bridgewater foi

para o Minnesota Vikings como a última escolha da primeira rodada (BONESTEEL, 2016).

Manziel foi um desastre, enquanto que Bridgewater tem se mostrando o prospecto indicado pelo estudo, sendo chamado para o *Pro Bowl* em 2016 e conduzindo o Vikings ao seu primeiro título de divisão desde 2009 (BONESTEEL, 2016). Uma possível explicação para a escolha do Browns é que quem encomendou a análise foi Joe Faixa que sairá de Cleveland antes do *Draft* tendo o novo CEO não seguido o estudo (THEPOSTGAME, 2014).

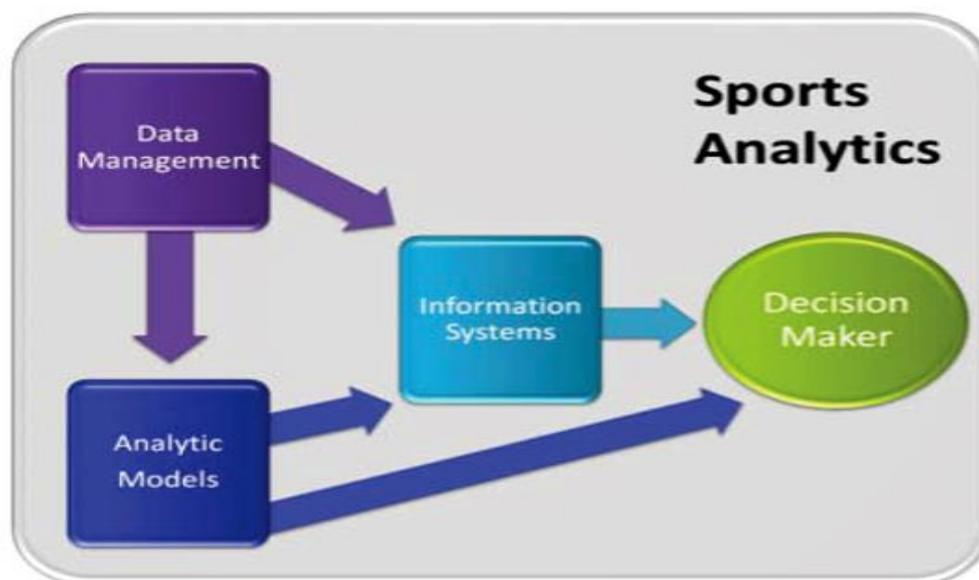
Não só os Browns cometeram erros, no *draft* de 2005 San Francisco 49ers seleciona Alex Smith, Nº 1 no primeiro round ao invés de Aaron Rogers, segundo Mike Nolan – ex-treinador do 49ers, a escolha foi baseada em características pessoais dos atletas, Rogers era arrogante enquanto Smith era uma escolha segura, um bom garoto. Nolan ainda comenta que pensaram que Smith seria melhor a longo prazo do que Rogers, o que não aconteceu Rogers deslanchou se tornando um dos melhores da liga enquanto Smith não conseguiu obter resultados (ORR, 2016).

Na temporada regular de 2015, entretanto, Alex Smith ficou em 8º lugar nas estatísticas do *Ranking* da NFL com um QB Rate de 84, jogando para o Kansas City Chiefs (NFL, 2016), mostrando que Smith ganhou vida sob o comando de Andy Reid, depois de uma década na NFL (ORR, 2016).

2.7. SPORTS ANALYTICS

Nos últimos anos o mundo viveu e vive a busca efusiva por informações, a habilidade de transformar dados brutos em informações complexas que traduzam um perfil de consumo ou prevejam um momento na história que possa gerar lucro e beneficiar pessoas e empresas, é de longe a busca incessante de qualquer gestor. No mundo esportivo não seria diferente, a informação trabalhada em modelos analíticos pode definir campeonatos, surge então o termo *Sports Analytics*. *Sports Analytics* é a gestão de dados históricos estruturados, a aplicação de modelos analíticos preditivos que utilizam esses dados e o uso de sistemas de informação para transcreve-los aos tomadores de decisão e assim permitir-lhes ajudar suas organizações a conquistar uma vantagem competitiva (ALAMAR E MEHROTRA, 2011).

Figura 8 – Sports Analytics Framework



Fonte: *Analytics*¹¹

O modelo de *framework* da Figura 8 mostra claramente o processo de geração de valor por meio da análise de dados, desde sua coleta até a tomada de decisão. *Data Management* (Gestão de dados) representa a aquisição e armazenagem de dados de forma estruturada em uma organização, sendo que no esporte esses dados são obtidos de várias fontes e podem ser apresentadas de muitas formas (ALAMAR E MEHROTRA, 2011).

Os modelos analíticos podem ser considerados o que de mais precioso se tem em um *framework*, uma vez que é a partir deles que os dados são transformados em informações que serão apresentadas a tomadores de decisão por meio de sistemas. A obtenção de modelos que traduzam os dados é uma luta exaustiva que exige a aplicação de ferramentas estatísticas de análise preditiva na obtenção de informações que prevejam o futuro (ALAMAR E MEHROTRA, 2011). Nos esportes isso pode envolver a projeção das carreiras profissionais dos jogadores amadores e identificar os pontos fortes e fracos dos adversários e da própria equipe (ALAMAR E MEHROTRA, 2011).

É exatamente sobre Modelos Analíticos voltados para a previsão de carreiras profissionais de jogadores amadores que este trabalho trata. Desde “*MoneyBall: The Art of Winning an Unfair Game*” (A Arte de Ganhar um Jogo Injusto) - que trata do uso de dados sobre atletas pelo time de beisebol Oakland Athletic’s, o mundo do beisebol não é

¹¹ Disponível em: <<http://analytics-magazine.org/beyond-moneyball-the-rapidly-evolving-world-of-sports-analytics-part-i/>>. Acessado em: 10 de julho 2016.

mesmo, os modelos matemáticos criados por Bill James e utilizados por Billy Beane e a forma como um time com poucos recursos consegue entrar para o livro dos recordes da liga americana de beisebol. Oakland Athletic's fez o que nenhum time havia feito nos últimos 102 anos de história da MLB: ganhou 20 vitórias consecutivas possuindo uma folha de pagamento de apenas US\$ 41.942.665, isso mudou a forma como o beisebol seleciona jogadores e realiza suas operações e transformou a história do esporte (LEWIS, 2004).

Os clubes de beisebol gastavam milhões em contratações de atletas, pois os jogadores podiam atuar como autônomos segundo uma lei aprovada na justiça americana, isso fez com que os salários dos atletas subissem de forma exagerada. Em quatro anos o salário médio das grandes ligas quase triplicará, passando de cerca de 52 mil dólares para quase 150 mil dólares por ano, o New York Yankees pagou 10 milhões de dólares pelo time inteiro em 1973 (LEWIS, 2004).

Segundo Lewis (2004) o Yankees chegou a pagar 126 milhões de dólares pelo seu elenco e muitos apostavam que o beisebol caminhava para uma disputa financeira e não mais nos jogos, ou seja, apostavam que quem possuíssem os jogadores mais bem pagos venceriam os jogos. Quando Alderson entrou no beisebol queria saber como o esporte funcionava, e foi o que fez. Chegou à conclusão de que tudo, desde as estratégias em campo até a avaliação dos jogadores, podia ser mais bem conduzido com uma investigação científica – hipóteses testadas pela análise de dados estatísticos históricos do beisebol – do que o saber coletivo da velha guarda do esporte (LEWIS, 2004, pg. 72).

Primeiro vieram os avanços drásticos na tecnologia de computação: reduziram-se expressivamente os custos de compilar e analisar a imensa quantidade de dados do beisebol. Então veio a explosão nos salários dos jogadores, o que aumentou expressivamente as vantagens de dispor de tais conhecimentos. 'Se vamos pagar 150 mil dólares por ano a esses caras, devíamos pelo menos saber até que ponto são bons' – o que significa saber o quanto produzem da defesa, e não só o quanto criam a cada ida ao bastão. Se esse argumento parecia convincente quando os jogadores de beisebol recebiam 150 mil dólares anuais, ficou cem vezes mais convincente quando passaram a receber 15 milhões por ano (LEWIS, 2004, pg. 88)".

Após Billy Beane sacudir o mundo do beisebol e *Moneyball* se tornar um sucesso nos Estados Unidos, no topo dos livros mais vendidos, iniciou-se uma revolução de estatísticas e tecnologias auxiliando o esporte, o beisebol começava a utilizar *Sports Analytics* e contrapor o culto da velha guarda. Os proprietários dos clubes começavam a questionar a abordagem que seguiam, assim os Cardinals de Bill DeWitt Jr, que acreditava

que a equipe não fora construída para durar contratou Luhnow para desenvolver um sistema para contratação de jogadores juntamente com um engenheiro da NASA, Sig Mejdal. Juntos levaram os Cardinals a *World Series* de 2013 onde 16 dos 25 jogadores tinham sido selecionados por eles (GREEN, 2016).

Jim Crane compra o Houston Astros em 2011 e chama Luhnow para ser gerente geral do clube, estando disposto a ir onde a análise de dados o levar. Os Astros tinham tido seu pior desempenho da história na temporada, e Crane queria transformá-los em um St. Louis Cardinals (GREEN, 2016). Em 2013 os Astros começaram a temporada com a menor folha de pagamento, USD 27 milhões e terminaram a temporada com apenas USD 13 milhões, sendo a pior equipe desde 1962-64 (GREEN, 2016).

Depois do *Moneyball* o beisebol evoluiu de tal forma que Luhnow criará um sistema que rastreará a bola em três dimensões e conseguia medir a velocidade, localização, aceleração movimento e eixo de rotação, possuindo segundo Mejdal uma capacidade preditiva inegável (GREEN, 2016). Uma matéria no *Wall Street Journal* chegou a comunicar que uma equipe de beisebol havia comprado um supercomputador para realizar simulações de mercado assim como fundos de *Hedge*¹²

Se espera que o mercado de *Sports Analytics* cresça de 2016 à 2021 de USD 123,7 milhões para USD 616,7 milhões, em uma taxa de crescimento composta de 37,9% (WOOD, 2016). O mercado de estatísticas voltadas ao esporte auxilia a gestão de várias organizações esportivas na análise de dados e na estimação de modelos preditivos que melhorem as decisões desde a seleção de jogadores até os jogos das temporadas.

Assim WOOD (2016) descreve que o mercado de *Sports Analytics* foi segmentado em regiões: América do Norte, Ásia-Pacífico, Europa, América Latina, Oriente Médio e África. A maior fatia do mercado está disposta nas regiões da América do Norte, Europa e Ásia-Pacífico em 2016, e estima-se que creça 38,4% de 2016 a 2021. América do Norte está rapidamente implementando a solução *Sports Analytics* devido ao ambiente de mercado dinâmico e maior adoção de tecnologias. América Latina, Oriente Médio e África estão testemunhando um crescimento lento na adoção da solução de *Sports Analytics* devido as baixas implementações de tecnologias e menos inclinação dos países do Oriente Médio para os esportes (WOOD, 2016).

¹² *Hedge* significa proteção. As operações de hedge são realizadas por empresas e investidores que desejam se proteger dos riscos das oscilações de preços, realizando fixações de preços no mercado. (CARVALHO, 2012)

3. ASPECTOS METODOLÓGICOS

Este item visa identificar os aspectos metodológicos que envolveram a pesquisa, identificando os instrumentos que foram utilizadas para a coleta dos dados, as variáveis dependentes e independentes utilizadas e os métodos e modelos de análise de dados escolhidos com base científica.

3.1. DEFINIÇÃO DAS VARIÁVEIS

Uma variável estatística é uma combinação linear de variáveis com pesos empiricamente determinados (HAIR et al., 2009) e seu objetivo é capturar o caráter multivariado da análise. Sendo que uma variável estatística com n variáveis ponderadas pode ser enunciada matematicamente como:

$$w_1X_1 + w_2X_2 + w_3X_3 + \dots + w_nX_n$$

Variáveis independentes (múltiplas) e variável dependente (única).

O rating do quarterback da NFL é calculado por uma fórmula e sua escala varia de 0 a 158,3, baseado na porcentagem de passes completos, nas jardas ganhadas por passe tentado, touchdowns por passes tentados e número de intercepções por passes tentados (PRO FOOTBALL, 2019), Miari (2017) identifica o *rating* conforme a fórmula a seguir:

$$\text{NFL Rating} = \left(\frac{\left(\frac{\text{COMP}}{\text{ATT}} - 0.3 \right) * 5 + \left(\frac{\text{YARDS}}{\text{ATT}} - 3 \right) * 0.25 + \left(\frac{\text{TD}}{\text{ATT}} \right) * 20 + 2.375 - \left(\frac{\text{INT}}{\text{ATT}} * 25 \right)}{6} \right) * 100$$

As variáveis selecionadas para tentar prever as habilidades futuras dos prospectos na NFL, medidas pelo *rating*, foram os dados históricos dos *quarterbacks* no college (ensino médio) e as métricas de suas habilidades físicas e psicológicas analisadas no combine, tendo estes jogado pelo menos 4 jogos no profissional e possuindo pelo menos 3 anos de experiência na liga desde 1987. Das variáveis do combine: *Height (in)*, *Weight (lbs)*, *Wonderlic test*, *40 yard dash*, *bench press*, *vert leap (in)*, *broad jump (in)*, *20 yard shuttle e 3 cone drill*, e as variáveis escolhidas do *college* foram: número de jogos (G), passes completos (comp), passes tentados (att), jardas (yds), *touchdowns* (TD), intercepções (INT) e o *rating* (RAT), sendo estas explicadas na Tabela 1.

Tabela 1 – Variáveis Dependentes e Independentes

Tipo	Variáveis	Medida	Descrição
Dependente	NFL_Rate	Escala Contínua	Rating da NFL
Independentes	Comb_Height (in)	Escala Contínua	Altura do prospecto em polegadas
	Comb_Weight (lbs)	Escala Contínua	Peso do prospecto em libras
	Comb_Wonderlic	Escala Contínua	Teste de inteligência (QI)
	Comb_40 Yard	Escala Contínua	Corrida de 40 jardas
	Comb_Bench Press	Escala Contínua	Número de repetições do supino reto
	Comb_Vert Leap (in)	Escala Contínua	Salto parado na vertical
	Comb_Broad Jump (in)	Escala Contínua	Salto em distância sem corrida
	Comb_Shuttle	Escala Contínua	Corrida de 20 jardas (ida e volta)
	Comb_3Cone	Escala Contínua	Percurso de corrida em "L"
	College_G	Escala Contínua	Quantidade de Jogos
	College_Comp	Escala Contínua	Número de Passes completos
	College_Att	Escala Contínua	Número de Passes Tentados
	College_Yds	Escala Contínua	Quantidade de Jardas percorridas
	College_TD	Escala Contínua	Número de Touchdowns
	College_INT	Escala Contínua	Número de Interceptações
	College_RAT	Escala Contínua	Rating College

Fonte: Elaboração própria

Todas as essas variáveis são classificadas como métricas ou contínuas de razão e não foram criadas variáveis adicionais. Dados métricos são utilizados quando indivíduos diferem em quantia ou grau em relação a um atributo em particular. As duas escalas de medida métrica são as intervalares e de razão (HAIR et al., 2009).

Segundo Hair et al. (2009) as escalas de razão representam a mais elevada forma de precisão de medida, pois possuem as vantagens de todas as escalas inferiores somadas à existência de um ponto zero absoluto.

3.2. TÉCNICA E INSTRUMENTO DE COLETA

Segundo Gerhardt (2009), pesquisa é definida como o procedimento racional e sistemático que tem como objetivo proporcionar respostas aos problemas que são propostos. A pesquisa desenvolve-se por um processo constituído de várias fases, desde a formulação do problema até a apresentação e discussão dos resultados (apud GIL, 2007 p. 17).

Deste modo, esta pesquisa utiliza referências constiuidas por material já elaborado de fontes diversas e dispersas (GERHARDT, 2009, apud FONSECA, 2002).

Os dados foram coletados de diferentes fontes secundárias na internet. Dados do Combine podem ser acessados no NFL Combine results, que é uma fonte facilmente acessível dos resultados dos testes, tanto por ano quanto por posição do jogador. Os dados são compilados e verificados manualmente de diferentes fontes da web incluindo NFL.com, CBS.com, ESPN.com, dentre outras, mas não garantem precisão em seus dados, sendo que as pontuações do teste Wonderlic e do 40 yard dash são sempre não oficiais. Dados históricos do college e da NFL foram coletados no Sports Reference, os dados do college antes dos anos 2000 é um produto do trabalho da Sports Data Research LLC e todos os dados posteriores a 2000 são da própria Sports Reference research, já os dados da NFL são, na maioria resultados do trabalho de Pete Palmer, Ken Pullis, and Gary Gillette.

Com a base de dados pronta para ser analisada, foi necessário escolher o melhor modelo dentre os vários disponíveis, baseando-se na relação de dependência a ser examinada, o número de variáveis dependentes e independentes (Tabela 1) e a escala de medida (HAIR et al., 2009).

3.3. DEFINIÇÃO DA TÉCNICA DE ANÁLISE DE DADOS

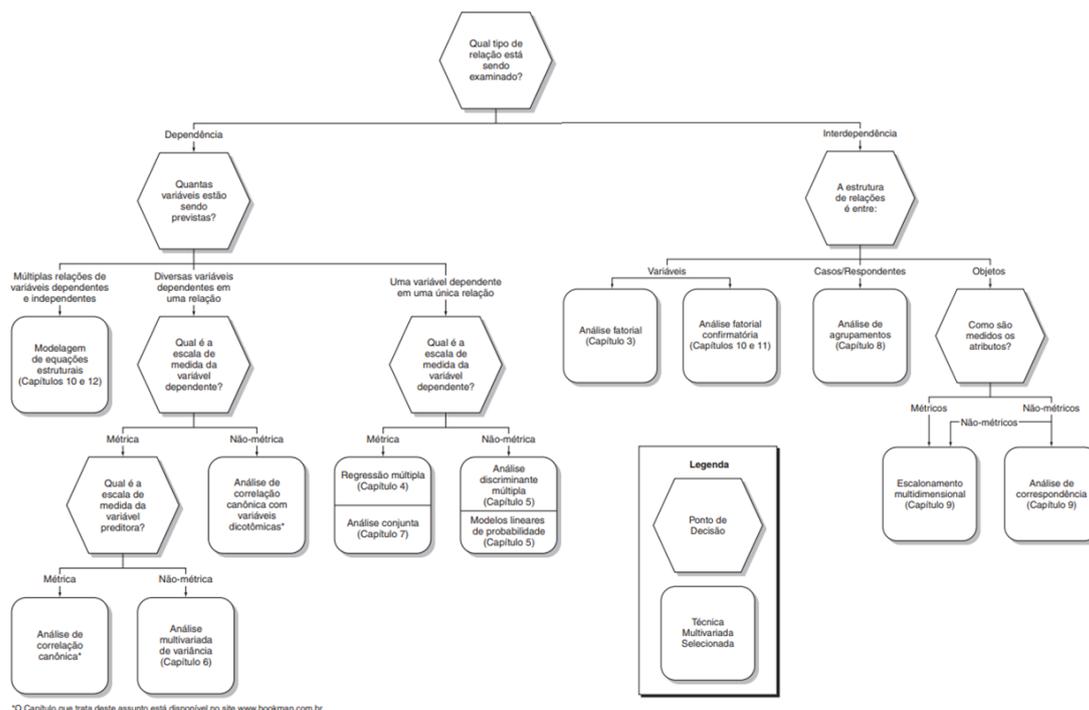
Este estudo é caracterizado por ser Aplicado, Quantitativo e Descritivo.

A pesquisa aplicada, segundo Gerhardt (2009), tem o objetivo de gerar conhecimentos para a aplicação prática, dirigidos à solução de problemas específicos. Esta pesquisa é exploratória pois visa proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícitos ou construir hipóteses (GERHARDT 2009), quantitativa porque seu resultado pode ser quantificado e descritiva visando descrever os fatos e fenômenos de determinada realidade (GERHARDT 2009).

A técnica escolhida para analisar os dados foi a análise multivariada. Segundo Hair et al. (2009), Análise Multivariada se refere a todas as técnicas estatísticas que simultaneamente analisam múltiplas medidas sobre indivíduos ou objetos sob investigação, tendo como objetivo medir, explicar e prever o grau de análise entre variáveis estatísticas.

Determinou-se então pelo diagrama (Figura 9) proposto por Hair et al. (2009) que a técnica a ser utilizada seria a de Regressão Múltipla, sendo a dependência entre a variáveis métricas mencionadas na tabela 1, o que queremos analisar.

Figura 9 – Seleção de uma técnica Multivariada



Fonte: HAIR et al. (2009)

A análise de regressão múltipla é uma técnica estatística que pode ser usada para analisar a relação entre uma variável dependente (critério) e várias variáveis independentes (preditoras) e tem como objetivo utilizar das variáveis preditoras, cujos valores são conhecidos, para prever os valores da variável dependente (HAIR et al., 2009).

Equação do modelo de regressão múltipla

$$Y = b_0 + b_1V_1 + b_2V_2 + \dots + b_nV_n e$$

Onde:

Y = Variável Estatística

b0 = Intercepto

b1, b2, bn = Coeficiente de regressão

V_1, V_2, V_n = Variáveis independentes

e = erro de previsão (resíduo)

Coefficiente de regressão é a variação estimada na variável dependente por variação unitária da variável independente (HAIR et al., 2009). Se o coeficiente de regressão é percebido como estatisticamente significativo (ou seja, o coeficiente é significativamente diferente de zero), o valor do coeficiente de regressão indica a extensão na qual a variável independente se associa com a dependente (HAIR et al., 2009). O intercepto tem valor explanatório apenas dentro do domínio de valores para as variáveis independentes e o erro de previsão, ou seja, a diferença entre os valores reais e previstos da variável dependente, é chamado de resíduo (HAIR et al., 2009).

4. DESCRIÇÃO E ANÁLISE DOS DADOS

Este capítulo apresenta as análises preliminares e pressupostos utilizados na regressão múltipla, bem como a parte teórica de estatística multivariada e seus resultados aplicados no objetivo deste estudo.

4.1. ANÁLISES PRELIMINARES

Após a coleta dos dados chegamos a uma base de dados com 616 *quarterbacks* de 1987 a 2019 que reunia informações da carreira do atleta no *college*, do resultado dos testes físicos e psicológicos do combine, o *round* e a posição do atleta no *draft* e também seus dados de carreira profissional na NFL.

Com esse resultado foram impostas duas premissas básicas, foi julgado que aqueles atletas com menos de quatro jogos e aqueles com menos de quatro anos na liga não participariam da análise, uma vez que queremos discutir a relação das variáveis que influenciam na seleção desses jogadores para representarem uma equipe na NFL ou seja, identificar quais as variáveis devem ser analisadas para tentar selecionar um futuro *quarterback* de sucesso.

Desta forma também os dados do draft não foram utilizados pois não são variáveis que representam a habilidade do prospecto em ser um bom *quarterback* no futuro.

Assim foi retirado os prospectos que não obtiveram sucesso ao ingressar na liga e aqueles que não participaram de pelo menos três jogos sobrando uma base de dados com 243 casos, a posteriores, retirando aqueles com menos de três anos, 222 casos, sendo estes o objeto de estudo desta pesquisa.

Segundo Pallant (2011) as estatísticas descritivas possuem inumeros usos incluindo descrever as características da amostra e verificar a violação de pressupostos subjacentes à técnica estatística utilizada na pesquisa.

4.2. PODER ESTATÍSTICO E TAMANHO DA AMOSTRA

Segundo Hair et al. (2009), o tamanho da amostra é o elemento mais influente sob o controle do pesquisador no planejamento da análise e seus efeitos são vistos diretamente no poder estatísticos do teste de significância e na generalização do resultado, sendo que amostras grandes tornam os testes de significância muito sensíveis.

O poder estatístico em regressão múltipla se refere a probabilidade de detectar-se como estatisticamente significativo um nível específico de R^2 ou um coeficiente de regressão em um nível de significância especificado para um dado tamanho de amostra (HAIR et al., 2009).

O tamanho da amostra também afeta a generalização dos resultados pela proporção entre observações e variáveis independentes, sendo uma regra geral que a razão jamais deve ficar abaixo de 5 para 1, ou seja, deve haver cinco observações para cada variável independente na variável estatística, sendo esse um nível mínimo para a modelagem de dados e caso essa proporção fique abaixo disso o pesquisador corre o risco de super ajustar a variável estatística à amostra, perdendo a generalização (HAIR et al., 2009). Um nível desejado está entre 15 e 20 observações, no entanto se um procedimento *stepwise* é empregado, é recomendado um nível de 50 para 1 (HAIR et al., 2009).

Pallant (2011) ainda cita que diferentes autores tendem a dar diversas diretrizes sobre o número de casos necessários para a regressão múltipla, recomendando cerca de 15 observações por preditor (apud Stevens, 1996) e ainda fornece uma fórmula para o cálculo dos requisitos de tamanho amostral levando em consideração as variáveis independentes (apud TABACHNICK E FIDELL, 2007)

$$N > 50 + 8m, \text{ onde}$$

N = Número de observações desejadas

M = Número de variáveis independentes

O grau de generalização é representado pelos graus de liberdade, calculados como:

Graus de liberdade (df) = Tamanho da amostra – Número de parâmetros estimados

ou

Graus de liberdade (df) = N – (Número de variáveis independentes + 1)

Quanto mais graus de liberdade, mais generalizáveis são os resultados. Graus de liberdade aumentam em uma dada amostra reduzindo-se o número de variáveis

independentes. Assim, a meta é conseguir a melhor precisão preditiva com o máximo de graus de liberdade (HAIR et al., 2009).

Nesta pesquisa chegamos a um total de 222 observações (como já mencionado no item 4.1) e possuímos um total de 16 variáveis independentes. Utilizando a fórmula mencionada por Pallant (2011) obtemos um número (N) de observações desejável por variável preditora de 178, porém analisando a tabela de análise descritiva (Tabela 2) verifica-se que seis variáveis não atingiram o número de observações desejáveis, sendo elas o teste Comb_wonderlic, Comb_Bench Press, Comb_Vert Leap, Comb_Broad Jump, Comb_Shuttle, e Comb_3Cone, sendo assim a decisão foi de utilizar o número mínimo de observações por variável, ou seja, 5 por 1, chegando a um nível desejável de 80 observações por variável preditora, retirando da análise apenas o Comb Bench Press com apenas 10 casos, no nível desejável de 15 para 1 o tamanho da amostra não atenderia a premissa de generalização.

Tabela 2 – Estatística Descritiva

	Nº	Mínimo	Máximo	Média	Erro Desvio
Comb_Height (in)	222,00	70,63	79,90	74,89	1,63
Comb_Weight (lbs)	222,00	179,00	265,00	219,91	14,04
Comb_Wonderlic	122,00	10,00	48,00	27,25	6,78
Comb_40 Yard	200,00	4,33	5,37	4,85	,19
Comb_Bench Press	10,00	3,00	24,00	14,60	7,32
Comb_Vert Leap (in)	176,00	22,00	39,00	31,27	3,29
Comb_Broad Jump (in)	173,00	93,00	127,00	109,52	6,91
Comb_Shuttle	169,00	3,87	4,87	4,31	,17
Comb_3Cone	114,00	6,66	7,92	7,17	,27
College_G	202,00	2,00	59,00	38,61	9,26
College_Comp	202,00	,00	1546,00	594,40	266,73
College_Att	202,00	1,00	2229,00	989,69	409,15
College_Yds	202,00	,00	19217,00	7628,64	3185,49
College_TD	202,00	,00	155,00	55,86	27,51
College_INT	202,00	,00	65,00	29,31	12,62
College_RAT	202,00	-200,00	178,20	134,49	29,00
NFL_Rate	220,00	,60	152,10	72,57	16,57
N válido (de lista)	1,00				

Fonte: Elaboração própria

O grau de generalização, representado pelos graus de liberdade foi de 206, já excluindo a variável Comb_Bench Press.

4.3. PRESSUPOSTOS DA ANÁLISE DE REGRESSÃO MÚLTIPLA

Alguns pressupostos são necessários no que envolve a regressão múltipla para verificar se os erros na previsão são um resultado de uma ausência real de relação entre as variáveis, ou eles são causados por alguma característica dos dados não acomodados pelo modelo (HAIR et al., 2009).

Assim os pressupostos a serem observados são:

- Multicolinearidade;
- Ausência de Outliers;
- Normalidade da distribuição dos termos de erro;
- Linearidade do fenômeno medido;
- Variância constante dos termos de erro (heterocedasticidade);
- Independência dos termos de erro.

4.3.1. MULTICOLINEARIDADE

Uma questão-chave na interpretação da variável estatística de regressão é a correlação entre as variáveis independentes. Esse é um problema de dados, e não de especificação de modelo. A situação ideal para um pesquisador seria ter diversas variáveis independentes altamente correlacionadas com a variável dependente, mas com pouca correlação entre elas próprias (HAIR et al., 2009)

Segundo Hair et al. (2009) a Multicolinearidade é o efeito combinado de duas ou mais variáveis independentes, sendo que a maneira mais simples de identificar colinearidade é um exame da matriz de correlação para as variáveis independentes. A presença de elevadas correlações (geralmente 0,90 ou maiores) é a primeira indicação de colinearidade substancial. No entanto, a falta de valores elevados de correlação não garante ausência de colinearidade. Colinearidade pode ser proveniente do efeito combinado de duas ou mais variáveis (HAIR et al., 2009).

Pallant (2011) propõe que não seja incluída variáveis com correlação superior a 0,7. Existem dois diagnósticos de colinearidade no processo de regressão múltipla, detectando problemas de multicolinearidade, sendo a tolerância e o VIF.

A tolerância é um indicador de quanto da variabilidade da variável independente selecionada não é explicada pelas outras variáveis independentes e é calcula pela fórmula $1 - R^2$. Caso o resultado da equação seja muito pequeno (inferior a 0,1) significa que existe múltipla correlação com outras variáveis sugerindo a possibilidade de multicolinearidade (PALLANT, 2011). O VIF (Variance inflation factor) é o inverso do valor da tolerância (1 dividido pela tolerância), valores acima de 10 indicaria multicolinearidade (PALLANT, 2011).

Analisando então a tabela de correlações (Tabela 3) percebemos que as variáveis salto vertical (Comb_Vert Leap), passes completos (College_Comp), passes tentados (College_att), jardas (College_Yds) e número de passes para touchdowns (College_TD) possuem uma correlação superior a 0,7, além disso seus valores de tolerância deram inferiores a 0,1 e os valores de VIF superiores a 10, com exceção do salto vertical, indicando a possibilidade de multicolinearidade (Tabela 4), o modelo completo pode ser observado no Apêndice A.

Tabela 3 - Correlações

		Correlações															
		NFL_Rate	Comb_Height (in)	Comb_Weight (lbs)	Com_Wonderlic	Comb_40 Yard	Comb_Vert Leap (in)	Comb_Broad Jump (in)	Comb_Shuttle	Comb_3Cone	College_G	College_Comp	College_Att	College_Yds	College_TD	College_INT	College_RAT
Correlação de Pearson	NFL_Rate	1,00	,05	,05	,22	-,02	,02	,14	-,03	-,11	,16	,24	,20	,21	,21	-,02	,18
	Comb_Height (in)	,05	1,00	,62	,17	,25	-,15	-,05	,17	,05	-,07	-,04	-,06	-,07	-,09	-,09	-,04
	Comb_Weight (lbs)	,05	,62	1,00	,09	,01	,10	,11	,00	-,10	-,07	,02	,00	,00	,02	-,19	,02
	Com_Wonderlic	,22	,17	,09	1,00	,20	-,19	-,15	-,12	-,25	,11	,13	,11	,09	,07	,02	,04
	Comb_40 Yard	-,02	,25	,01	,20	1,00	-,73	-,67	,46	,44	,01	,09	,10	,05	,00	,20	-,11
	Comb_Vert Leap (in)	,02	-,15	,10	-,19	-,73	1,00	,75	-,48	-,35	-,05	,01	-,01	,03	,06	-,11	,03
	Comb_Broad Jump (in)	,14	-,05	,11	-,15	-,67	,75	1,00	-,45	-,47	-,06	-,03	-,04	-,02	-,01	-,17	,09
	Comb_Shuttle	-,03	,17	,00	-,12	,46	-,48	-,45	1,00	,61	-,01	,00	,01	-,02	-,05	,12	-,19
	Comb_3Cone	-,11	,05	-,10	-,25	,44	-,35	-,47	,61	1,00	-,06	-,04	,00	-,02	-,01	,13	-,22
	College_G	,16	-,07	-,07	,11	,01	-,05	-,06	-,01	-,06	1,00	,62	,65	,64	,56	,51	,34
	College_Comp	,24	-,04	,02	,13	,09	,01	-,03	,00	-,04	,62	1,00	,98	,98	,90	,67	,31
	College_Att	,20	-,06	,00	,11	,10	-,01	-,04	,01	,00	,65	,98	1,00	,97	,87	,76	,25
	College_Yds	,21	-,07	,00	,09	,05	,03	-,02	-,02	-,02	,64	,98	,97	1,00	,93	,68	,35
	College_TD	,21	-,09	,02	,07	,00	,06	-,01	-,05	-,01	,56	,90	,87	,93	1,00	,52	,42
	College_INT	-,02	-,09	-,19	,02	,20	-,11	-,17	,12	,13	,51	,67	,76	,68	,52	1,00	,04
	College_RAT	,18	-,04	,02	,04	-,11	,03	,09	-,19	-,22	,34	,31	,25	,35	,42	,04	1,00

Fonte: Elaboração própria

Pallant (2011) sugere a remoção das variáveis independentes altamente correlacionadas do modelo, desta forma testamos o modelo removendo cada variável e todas as combinações apresentaram valores de significância acima de 0,05 e multicolinearidade entre passes completos (College_Comp), tentados (College_Att), número de jardas percorrida (College_Yds) e número de passes para *touchdowns* (College_TD), sendo assim iremos remover as variáveis College_comp, College_Att e College_Yds, considerando apenas College_TD que influência diretamente na vitória dos jogos. Optou-se por não remover o salto vertical por não apresentar valores de tolerância e de VIF discrepantes, resultando em um total de 12 variáveis independentes, alterando a premissa de tamanho da amostra para pelo menos 60 observações, não influenciando nas variáveis independentes utilizadas.

Tabela 4 – Estatísticas de Colinearidade

		Descrição	Tolerância	VIF
1	(Constante)			
	Comb_Height (in)	Altura do prospecto em polegadas	,52	1,94
	Comb_Weight (lbs)	Peso do prospecto em libras	,52	1,93
	Com_Wonderlic	Teste de inteligência (QI)	,75	1,33
	Comb_40 Yard	Corrida de 40 jardas	,35	2,88
	Comb_Vert Leap (in)	Salto parado na vertical	,29	3,40
	Comb_Broad Jump (in)	Salto em distância sem corrida	,34	2,94
	Comb_Shuttle	Corrida de 20 jardas (ida e volta)	,51	1,96
	Comb_3Cone	Percurso de corrida em "L"	,44	2,25
	College_G	Quantidade de Jogos	,50	1,99
	College_Comp	Número de Passes completos	,02	57,01
	College_Att	Número de Passes Tentados	,01	66,80
	College_Yds	Quantidade de Jardas percorridas	,02	47,44
	College_TD	Número de Touchdowns	,09	10,67
	College_INT	Número de Interceptações	,24	4,20
	College_RAT	Rating do College	,63	1,59

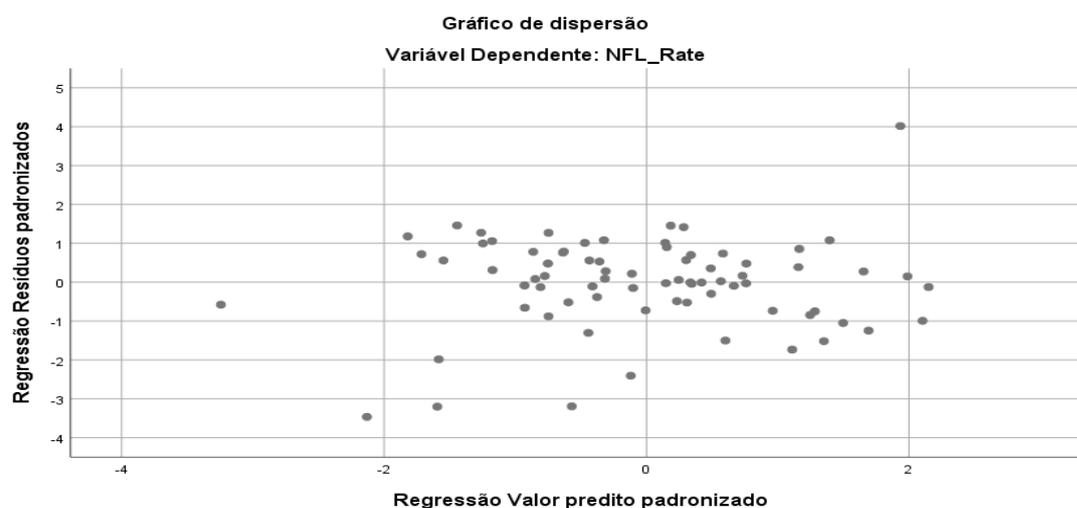
Fonte: Elaboração própria.

Após remover as variáveis e refazer o modelo, não foi verificado indícios de multicolinearidade por meio dos métodos já apresentadas neste item (Apêndice B).

4.3.2. OUTLIERS

Os outliers podem ser verificados inspecionando o gráfico de dispersão dos resíduos padronizados, os resíduos devem estar distribuídos de maneira retangular com mais pontos concentrados no centro, mais próximo do ponto 0. Desvios sugerem alguma violação de premissas, e Pallant apud Tabachnick e Fidell (2007) define outliers como casos que possuem um resíduo padronizado superior a 3,3 ou inferior a -3,3.

Gráfico 1 – Dispersão dos Dados



Fonte: Elaboração própria.

Para verificar se os outliers estão exercendo influência indevida nos resultados para o modelo pode-se verificar a Distância de Cook na tabela de Estatísticas de Resíduos, casos com valores maiores que 1 são um problema em potencial (PALLAND *apud* TABACHNICK E FIDELL, 2007).

Tabela 5 – Estatísticas de Resíduos

	Mínimo	Máximo	Média	Erro Desvio	N
Valor previsto	48,93	88,24	72,35	7,85	76,00
Erro Valor previsto	-3,24	2,15	-,03	1,08	76,00
Erro padrão do valor previsto	4,25	22,52	6,74	2,21	76,00
Valor previsto ajustado	51,41	94,00	72,68	8,68	74,00
Resíduo	-56,44	65,43	-,33	19,01	75,00
Erro Resíduo	-3,47	4,02	-,02	1,17	75,00
Resíduos Resíduo	-3,66	4,39	-,01	1,29	74,00

de Estud.	-64,92	78,15	-,18	23,01	74,00
Resíduos de Estud.	-4,01	5,05	-,02	1,36	74,00
Mahal. Distância	5,22	173,07	16,26	19,11	76,00
Distância de Cook	,00	,23	,02	,04	74,00
Valor de ponto alavanca centralizado	,06	1,90	,18	,21	76,00
a. Variável Dependente: NFL_Rate					

Fonte: Elaboração própria.

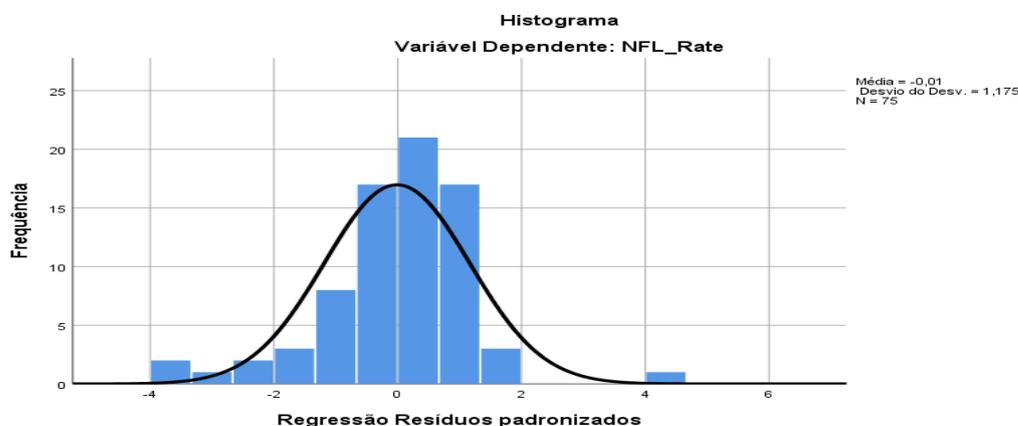
Pallant (2011) ainda menciona ser comum entrar vários resíduos periféricos, e encoraja que se o pesquisador encontrar apenas alguns, não precisará tomar nenhuma ação. Analisando o gráfico de dispersão (Gráfico 1) verificamos a existência de apenas 4 observações que violaram as premissas e como os valores da Distância de Cook não estão superiores a 1, sendo o valor máximo de Cook 0,23, sugere-se que não existem problemas em manter as observações.

4.3.3. NORMALIDADE

Talvez a violação de suposição mais frequentemente encontrada seja a não-normalidade das variáveis independentes ou dependente ou ambas. O diagnóstico mais simples para o conjunto de variáveis independentes na equação é um histograma de resíduos, com uma verificação visual para uma distribuição que se aproxima da normal (HAIR et al. 2009).

Ao verificar o histograma vemos que os resíduos padronizados do modelo parecem seguir uma distribuição normal, não violando esta premissa.

Gráfico 2 – Distribuição normal

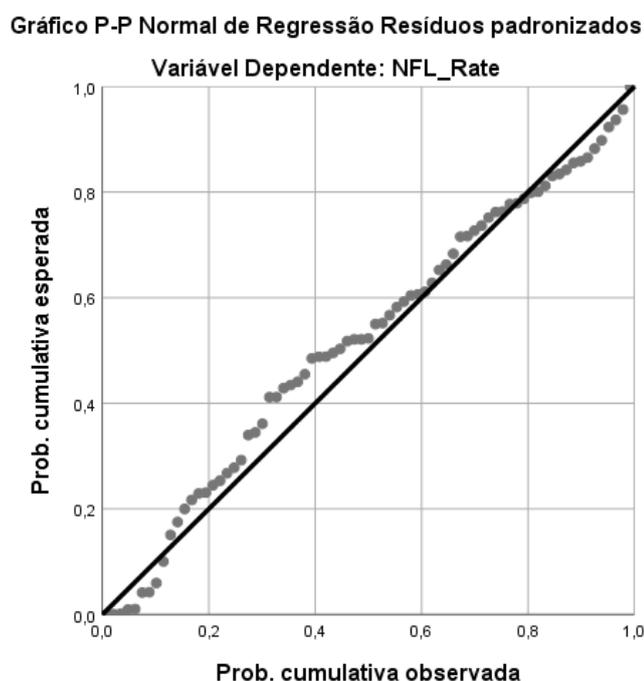


Fonte: Elaboração própria.

4.3.4. LINEARIDADE

A linearidade da relação entre variáveis dependente e independentes representa o grau em que a variação na variável dependente é associada com a variável independente (HAIR et al. 2009). Hair (2009) ainda menciona que a linearidade de qualquer relação bivariada é facilmente examinada por meio de gráfico de resíduos, e que qualquer padrão curvilíneo consistente indica que uma ação corretiva aumentará a precisão do modelo bem como a validade dos coeficientes estimados.

Gráfico 3 – Distribuição normal dos Resíduos padronizados



Fonte: Elaboração própria.

Verificando então o gráfico de distribuição normal dos resíduos padronizados (Gráfico 3) podemos analisar que os resíduos demonstram uma distribuição normal, não sendo necessário uma transformação nos dados.

4.3.5. HETEROCEDASTICIDADE

A presença de variâncias desiguais (heteroscedasticidade) é uma das violações mais comuns de suposições. O diagnóstico é feito com gráficos de resíduos ou testes

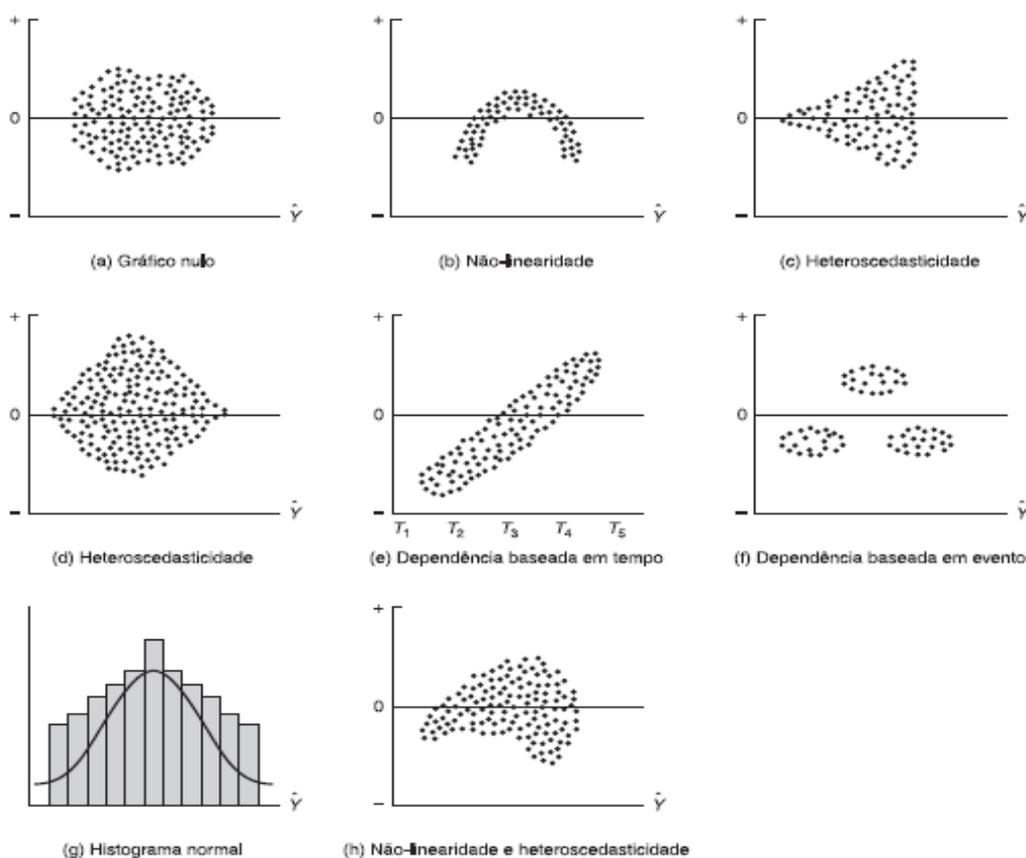
estatísticos simples. A representação gráfica de resíduos (estudentizados) versus os valores dependentes previstos e a sua comparação com o gráfico nulo (Figura 10) (HAIR et al. 2009).

Considera-se então o gráfico de dispersão (Gráfico 1) em comparação com o gráfico nulo exposto por Hair (2009) que os dados não apresentam heterocedasticidade, pelo padrão da distribuição dos resíduos.

4.3.6. INDEPENDÊNCIA

Assumimos em regressão que cada valor previsto é independente, o que significa que o valor previsto não está relacionado com qualquer outra previsão, ou seja, eles não são seqüenciados por qualquer variável. Podemos identificar melhor tal ocorrência fazendo o gráfico de resíduos em relação a qualquer variável seqüencial possível. Se os resíduos forem independentes, o padrão deverá parecer aleatório e semelhante ao gráfico nulo de resíduos (HAIR et al. 2009).

Figura 10 – Análise gráfica de resíduos



Fonte: Hair et al. 2009

Na mesma comparação da heterocedasticidade concluímos que os dados apresentam um padrão consistente analisado pelo gráfico de dispersão muito adequado segundo Hair (2009) com o gráfico nulo, subentende-se que os dados não possuem variância constante.

4.3.7. ESTIMAÇÃO DO MODELO E AVALIAÇÃO DO AJUSTE GERAL

Segundo Hair (2009) após ter especificado os objetivos da análise de regressão, selecionado as variáveis dependentes e independentes, abordado as questões de planejamento e pesquisa e avaliado se as variáveis atendem às suposições da regressão o modelo pode então ser estimado, sendo necessários alguns passos:

1. Selecionar um método para especificar o modelo de regressão a ser estimado;
2. Avaliar a significância estatística do modelo geral na previsão da variável dependente;
3. Determinar se alguma das observações exerce uma influência indevida nos resultados.

4.3.7.1. SELEÇÃO DA TÉCNICA DE ESTIMAÇÃO

Em alguns casos o conjunto de variáveis independentes pode ser especificado exatamente e o modelo de regressão é essencialmente usado em uma abordagem confirmatória, porém em outros casos, o pesquisador pode utilizar de métodos de busca sequencial ou combinatório (HAIR et al. 2009).

Hair (2009) dispõe que a abordagem mais simples é a confirmatória porém exige que o pesquisador especifique as variáveis independentes a serem incluídas, sendo este completamente responsável pelas comparações entre elas e a precisão do modelo. Já em métodos sequenciais, são acrescentadas ou eliminadas até que alguma medida de critério seja alcançado, maximizando a previsão ao mesmo tempo que emprega o menor número de variáveis (HAIR et al. 2009).

Para este modelo a abordagem escolhida foi a confirmatória uma vez que para utilizar-se de modelos sequenciais como o *stepwise* seria necessário mais observações, como mencionado no item 4.2, ferindo o princípio do tamanho da amostra.

4.3.7.2. AVALIAÇÃO DO MODELO

Para avaliar o modelo devemos analisar seu poder estatístico, como já mencionado no item 4.2. O R^2 indica o quanto da variação na variável dependente é explicado pelo modelo (PALLANT, 2011) e o R^2 Ajustado é a modificação do R^2 que considera o número de variáveis independentes incluídas na equação de regressão e o tamanho da amostra (HAIR et al. 2009). Neste caso, o valor de R^2 é 0,16 e o R^2 Ajustado é 0,03 como pode ser notado na Tabela 6, assim o modelo com as variáveis independentes utilizadas explica, juntas, 16 por cento da variação da variável dependente e considerando o modelo ajustado pela quantidade relevante de variáveis independentes e tamanho de amostra variados, o modelo explica apenas 3% da variabilidade da variável dependente em função das independentes, sendo este valor muito baixo.

Tabela 6 – Resumo do Modelo

Resumo do modelo^b									
Modelo	R ²	R ² ajustado	Erro padrão da estimativa	Mudança de R ²	Estatísticas de mudança			Sig. Mudança F	
					Mudança F	f1	f2		
	40 ^a	16	,03	16,29	,16	1,26	2,00	9,00	,26

a. Preditores: (Constante), College_RAT, Comb_Weight (lbs), Com_Wonderlic, College_INT, Comb_Vert Leap (in), Comb_3Cone, College_G, Comb_Height (in), Comb_Shuttle, College_TD, Comb_40 Yard, Comb_Broad Jump (in)

b. Variável Dependente: NFL_Rate

c. Fonte: Elaboração própria.

Para testar a hipótese de que a quantidade de variação explicada pelo modelo de regressão é maior que a previsão de referência (ou seja, que R^2 é significativamente maior que zero), a razão F é calculada (HAIR et al. 2009):

$$\text{Razão } F = \frac{\frac{\text{Soma de quadrados regressão}}{\text{Graus de liberdade regressão}}}{\frac{\text{Soma de quadrados residual}}{\text{Graus de liberdade residual}}} = \frac{\text{SS regressão}}{\text{df regressão}} \div \frac{\text{SS residual}}{\text{df residual}}$$

Onde:

df regressão = Número de coeficientes estimados (incluindo interceto) – 1

df regressão = Tamanho da amostra – Número de coeficientes

estimados (incluindo interceito)

Para medir a significancia do modelo é necessário verificar a tabela *ANOVA* que possui um teste de significancia *F*. O modelo não atingiu significância estatística, com poder de 0,26 (isso significa que $p > 0,05$) não se rejeita a hipótese nula, ou seja, o ajuste do modelo somente com o intercepto e o modelo são iguais.

Tabela 7 – Teste de significância

ANOVA ^a						
Modelo	Soma dos Quadrados	df	Quadrado Médio	Z	Sig.	
Regressão	4019,93	12,00	334,99	1,26	,26 ^b	
Resíduo	20973,77	79,00	265,49			
Total	24993,70	91,00				

a. Variável Dependente: NFL_Rate

b. Preditores: (Constante), College_RAT, Comb_Weight (lbs), Comb_Wonderlic, College_INT, Comb_Vert Leap (in), Comb_3Cone, College_G, Comb_Height (in), Comb_Shuttle, College_TD, Comb_40 Yard, Comb_Broad Jump (in)

c. Fonte: Elaboração própria.

4.3.8. AVALIANDO AS VARIÁVEIS INDEPENDENTES

O próximo passo é analisar quais as variáveis incluídas no modelo que contribuíram para a previsão da variável dependente (PALLANT, 2011). Para comparar a contribuição de cada variável é importante observar os coeficientes padronizados *Beta*, aqueles com maior contribuições terão valores mais elevados, desta forma a variável que teve a maior contribuição foi o salto em distância sem corrida (Comb_Broad Jump), com valor de *beta* de 0,34. Observando então os valores de significância (Sig.) conclui-se que nenhuma variável está tendo uma contribuição significativa para a previsão da variável dependente pois todos os valores são maiores que 0,05.

Tabela 8 - Coeficientes

Modelo	Coeficientes não padronizados		Coeficientes padronizados Beta	t	Sig.	Correlações		
	B	Erro Erro				Ordem zero	Parcial	Parte
(Constante)	-90,77	134,39		-,68	,50			
Comb_Height (in)	,17	1,46	,02	,12	,91	,05	,01	,01
Comb_Weight (lbs)	-,04	,17	-,03	-,21	,83	,05	-,02	-,02
Comb_Wonderlic	,54	,29	,22	1,87	,07	,22	,21	,19
Comb_40 Yard	4,27	15,16	,05	,28	,78	-,02	,03	,03
Comb_Vert Leap (in)	-,71	,95	-,14	-,75	,45	,02	-,08	-,08
Comb_Broad Jump (in)	,82	,42	,34	1,96	,05	,14	,22	,20
Comb_Shuttle	7,58	13,51	,08	,56	,58	-,03	,06	,06
Comb_3Cone	1,68	9,17	,03	,18	,85	-,11	,02	,02
College_G	,16	,24	,09	,66	,51	,16	,07	,07
College_TD	,14	,09	,23	1,56	,12	,21	,17	,16
College_INT	-,23	,18	-,18	-1,25	,21	-,02	-,14	-,13
College_RAT	,03	,07	,05	,40	,69	,18	,05	,04

Fonte: Elaboração própria.

5. CONSIDERAÇÕES FINAIS

A regressão múltipla padrão foi usada para avaliar quais variáveis são importantes serem analisadas no recrutamento e seleção de quarterbacks por um time da NFL. Análises preliminares foram realizadas para garantir a não violação das premissas de generalização, normalidade, *outliers*, linearidade, multicolinearidade e homoscedasticidade. Algumas variáveis foram então retiradas da pesquisa por apresentarem multicolinearidade, e todo o processo foi revisto, desta foram 12 variáveis independentes foram selecionadas para prever a variável dependente (*Rating* da NFL) resultando em 3% (R^2 Ajustado) da variação total explicada pelo modelo como razão F (12, 79) = 1,26 e com significância estatística de apenas 0,26. No modelo final, nenhuma das variáveis independentes foram estatisticamente significantes, com valores de beta baixos e significância maiores que 0,05, tornando o modelo impraticável na previsão do *rating*.

Verificado então que as variáveis independentes não são significantes para prever a variação da variável dependente, mostrando que as variáveis do Combine e do College não influenciam no desempenho do QB no futuro, propôs-se testar se elas influenciam na posição de escolha do prospecto, analisando então o modelo de outra perspectiva, a do jogador.

As análises preliminares não mostraram qualquer violação nas premissas da regressão múltipla, resultando em um modelo que explica 23% da variabilidade da variável dependente (R^2 ajustado = ,23), com razão F (12, 79) = 3,33 e com significância estatística. Avaliando cada uma das variáveis independentemente, aquelas que apresentaram uma significância abaixo de 0,05, em ordem de importância foram: Passes para touchdowns (-,35) e 3 Cone Drill (,33) as demais variáveis não criaram contribuições significativas para o modelo, conforme pode ser visto no Apêndice C.

5.1. CONTRIBUIÇÕES DA PESQUISA

Esta pesquisa contribui de forma eficaz na análise de recrutamento e seleção de QBs na NFL, uma vez que não identifica uma relação entre as variáveis individuais do atleta, divulgadas como as utilizadas para escolher e medir o desempenho dos jogadores na liga. Assim, conclui-se que com apenas as variáveis apresentadas como públicas – os testes no *Combine* e o histórico do jogador no *College* – não se pode inferir que o atleta

poderá ou não se tornar um bom jogador, sendo que estas não influenciam em seu desempenho na liga profissional, porém podem influenciar sua posição de escolha do *draft*, principalmente os passes para touchdowns e o resultado do teste de agilidade do combine 3 Cone Drill.

5.2. SUGESTÕES PARA PESQUISAS FUTURAS

Com os avanços das tecnologias de informação, de levantamento de dados e inteligência artificial, um estudo mais aprofundado pode ser realizado com variáveis comportamentais dos jogadores afins de aferir o objetivo deste estudo.

A pesquisa teve uma variável métrica como variável dependente, um estudo pode ser realizado considerando então uma variável não métrica como variável dependente, podendo assim partir para outros modelos de análise multivariada, como a Análise Discriminante e a Regressão logística, sendo talvez mais eficientes na adequação dos dados e obtendo uma resposta estatística mais significativa.

Pode-se também ser acrescentado a pesquisa variáveis do time como pontos, vitórias e eficiência dos Wide Receivers, podendo o desempenho do QB ser influenciado pela precisão das demais posições.

Outro estudo pode ser realizado clusterizando os atletas e verificando então, por grupo, quais as variáveis individuais e coletivas que podem influenciar no desempenho do time no meio pré-definido pelo pesquisador.

5.3. LIMITES DO ESTUDO

O estudo se limitou a variáveis de desempenho físico do atleta e no histórico de partidas anteriores na liga universitária americana. Uma hipótese seria acrescentar, se possível, variáveis comportamentais. Além disso, com a quantidade de dados que possuímos hoje não podemos utilizar métodos que levem em consideração parâmetros pressupostos, como por exemplo o método *stepwise*.

REFERÊNCIAS

ANDRADE, GUSTAVO. **Tudo sobre NBA: história, equipes, regras, draft e campeões**. Esportelândia, 29 de Julho de 2019. Disponível em: <https://www.esportelandia.com.br/basquete/nba/#Como_funciona_o_draft_da_NBA>. Acessado em 24 de Novembro de 2019

ARTHUR, Rachel. **The Marketing Power of Sports' Stars**. The New York Times, 04 de abril de 2016. Disponível em; <<https://www.nytimes.com/2016/04/05/fashion/sports-athletes-marketing.html>>. Acessado em: 24 de Novembro de 2019.

BELZER, JASON. **Thanks To Roger Goodell, NFL Revenues Projected To Surpass \$13 Billion In 2016**. Forbes, 29 de fevereiro de 2016. Disponível em: <<http://www.forbes.com/sites/jasonbelzer/2016/02/29/thanks-to-roger-goodell-nfl-revenues-projected-to-surpass-13-billion-in-2016/#6fbf27233278>>. Acessado em: 08 de Julho de 2016.

BONESTEEL, M. The Browns' \$100,000 study told them to take Teddy Bridgewater. They took Johnny Manziel. The Washington Post, 11 de Maio de 2016. Disponível em: <<https://www.washingtonpost.com/news/early-lead/wp/2016/03/11/the-browns-100000-study-told-them-to-take-teddy-bridgewater-they-took-johnny-manziel/>>. Acessado em: 19 de Julho de 2016.

CARVALHO, H. **Hedge: O que é e como fazer?**. Disponível em: <<http://hcinvestimentos.com/2012/06/25/hedge/>>. Acessado em 22 de Julho de 2016.

CARVALHO, IEDA MARIA VECCHIONI. **Recrutamento e seleção por competências**. Rio de janeiro: Editora FGV, 2008.

CRUYFF, Institute. **The athlete as a social media marketing product**. Publicado em 23 de Novembro de 2015. Disponível em: <<https://johancruyffinstitute.com/en/blog-en/the-athlete-as-a-social-media-marketing-product/>>. Acessado em: 24 de Novembro de 2019.

ECKSTEIN, Jaob. How The NFL Makes Money. Publicado na Investopedia em 24 de setembro de 2019. Disponível em: <<https://www.investopedia.com/articles/personal-finance/062515/how-nfl-makes-money.asp>>. Acessado em 30 de Novembro de 2019.

ESPN. **A História da NFL**. Disponível em: <<http://espn.uol.com.br/infografico/guiafutebolamericano/ahistoriadanfl/>>. Acessado em: 08 de Julho de 2016.

FORBES. **Sports Money: 2019 NFL Valuations**. Disponível em: <<https://www.forbes.com/nfl-valuations/list/#tab:overall>>. Acessado em: 30 de Novembro de 2019.

_____. **The World's Highest Athletes.** Disponível em: <<https://www.forbes.com/athletes/list/#tab:overall>>. Acessado em: 30 de Novembro de 2019.

FORLIN, F. **Sobre quarterbacks e sua real importância.** Publicado em 11 de Maio de 2013 em NFL S.A. Disponível em: <<https://nflsa.wordpress.com/2013/05/11/sobre-quarterbacks-e-sua-real-importancia/>>. Acessado em 22 de Julho de 2016.

GABRIEL, LUCAS. 4 Ps do Marketing entenda o conceito do Mix de Marketing. Rockcontent, 13 de Agosto de 2018. Disponível em: <<https://rockcontent.com/blog/4-ps-do-marketing/>>. Acessado em 24 de Novembro de 2019.

GERHARDT, Tatiana Engel; SILVEIRA, Denise Tolfo. **Métodos de pesquisa.** Porto Alegre: Ufrgs, 2009. 120 p.

GREEN, J. **Can a former McKinsey consultant build a championship baseball team? The houston Astros are finding out.** Publicado na Semana de Negócios da Bloomberg.

GUTERMAN, MARCOS. **O futebol explica o Brasil: Uma história da maior expressão popular do Brasil.** São Paulo: Editora Contexto, 2010.

HAIR, Joseph F. et al. **Análise multivariada de dados.** 6. ed. Porto Alegre: Bookman Companhia Editora Ltda, 2009. 688 p.

HENNINGER, Brian. **The History Of American Football.** Dynasty Sports Blog. Publicado em 11 de Janeiro de 2018. Disponível em: <<https://www.shopdynastysports.com/blogs/news/the-history-of-american-football>>. Acessado em 30 de Novembro de 2019.

HOW MUCH. **Which Professional Sports Leagues Make the Most Money?** Disponível em: <<http://howmuch.net/articles/sports-leagues-by-revenue>>. Acessado em: 08 de Julho de 2016.

KOTLER, Philip; KELLER, Kevin Lane. **ADMINISTRAÇÃO DMARKETING.** 14. ed. São Paulo: Pearson, 2013. 765 p.

LANE, Barnaby. Cristiano Ronaldo reportedly makes more money being an influencer on Instagram than he does playing soccer for Juventus. Bussiness Insider, 16 de Outubro de 2019. Disponível em: <<https://www.businessinsider.com/cristiano-ronaldo-makes-more-money-from-instagram-than-juventus-2019-10>>. Acessado em: 24 de Novembro de 2019.

LEWIS, MICHAEL. **Moneyball: O Homem que Mudou o Jogo.** Editora Intrinseca LTDA, Rio de Janeiro, 2004.

LILLY, BRANDON. **College football explained**. The Guardian, 10 de Outubro de 2012. Disponível em: <<https://www.theguardian.com/sport/blog/2012/oct/10/college-football-explained-ncaa#comments>>. Acessado em 09 de Julho de 2016.

MIARI, CAIO. **O que é o rating no futebol americano?**. Publicado em 09 de Junho de 2017 na Shotgun: Tudo sobre Futebol Americano. Disponível em: <<http://shotgun.football/o-que-e-o-rating-no-futebol-americano/>>. Acessado em: 17 de Novembro de 2019.

NCAA. **What is the NCAA?**. Disponível em <<http://www.ncaa.org/about/resources/media-center/ncaa-101/what-ncaa>>. Acessado em: 09 de Julho de 2016.

NFL SHOP. **Top selling jerseys**. Disponível em: <<https://www.nflshop.com/top-selling-jerseys/x-249520+z-98608802-3043109587>>. Acessado em: 24 de Novembro de 2019.

NFL UP. **Training for the NFL Combine bench press test**. Fevereiro de 2016. Disponível em: <<http://www.nfl.com/news/story/0ap3000000637311/article/training-for-the-nfl-combine-bench-press-test>>. Acessado em: 09 de Julho de 2016.

NFL WORKOUTS. **What Goes On At The Combine**. Disponível em: <<http://www.nfl.com/combine/workouts>>. Acessado em: 09 de Julho de 2016.

NFL, OPERATION. **The Rules of the Draft**. Disponível em: <<http://operations.nfl.com/the-players/the-nfl-draft/the-rules-of-the-draft/>>. Acessado em: 09 de Julho de 2016.

NFL, **Statistics**. Disponível em: <<http://www.nfl.com/stats/categorystats?archive=false&conference=null&statisticPositionCategory=QUARTERBACK&season=2015&seasonType=POST&experience=&tabSeq=1&qualified=false&Submit=Go>>. Acessado em: 21 de Julho de 2016.

NFL. **History**. Disponível em: <<http://www.nfl.com/history/chronology/1869-1910>>. Acessado em: 08 de Julho de 2016.

ORR, C. **Nolan explains taking Alex Smith over Aaron Rodgers**. Publicado em 26 de abril de 2016 na NFL. Disponível em: <<http://www.nfl.com/news/story/0ap3000000656167/article/nolan-explains-taking-alex-smith-over-aaron-rodgers>>. Acessado em: 21 de Julho de 2016.

OZANIAN, MIKE. **Super Bowl 50 Will Generate \$620 Million For NFL**. Forbes, 30 de janeiro de 2016. Disponível em: <<http://www.forbes.com/sites/mikeozanian/2016/01/30/super-bowl-50-will-generate-620-million-for-nfl/#2b40da1072dd>> . Acessado em: 08 de Julho de 2016.

PALLANT, Julie. **SPSS Survival Manual**. 4. ed. Sydney: Allen & Unwin, 2011. 345 p.

PAOLI, PB; SILVA, CD; SOARES, AJG. **Tendência atual da detecção, seleção e formação de talentos no futebol brasileiro.** Universidade Federal de Viçosa – MG/Brasil; Universidade Gama Filho – RJ/Brasil. Revista Brasileira de Futebol, 2008.

PRO FOOTBALL. **NFL Passer Rating Calculator.** Disponível em: <<https://www.pro-football-reference.com/about/qb-rating.htm>>. Acessado em: 17 de Dezembro de 2019.

ROBERTSON, MIKE. **The Ultimate Guide to the NFL Combine.** Robertson Training Systems, 23 de Fevereiro de 2016. Disponível em: <<http://robertsontrainingsystems.com/blog/ultimate-guide-nfl-combine/>>. Acessado em: 09 de Julho de 2016.

SALVADOR, Alexandre. **Efeito Neymar: PSG vende 1 milhão de euros em camisas em um dia,** Revista Veja, 04 de Agosto de 2017. Disponível em: <<https://veja.abril.com.br/esporte/efeito-neymar-psg-vende-1-milhao-de-euros-em-camisas-em-um-dia/>>. Acessado em: 24 de Novembro de 2019.

THEPOSTGAME STAFF. **Browns Disregard \$100,000 Study That Says Bridgewater Is Draft's Best QB.** Artigo digital publicado em 05 de Setembro de 2014. Disponível em: <<<http://www.thepostgame.com/blog/dish/201405/browns-100000-study-finds-teddy-bridgewater-best-qb-draft>>>. Acessado em: 19 de Julho de 2016>.

VEIGA, Vinícius. **Em Todos os detalhes, entenda como funciona o draft da MLB.** Casa do Beisebol, 06 de Junho de 2017. Disponível em: <<http://casadobeisebol.com.br/draftmlb/>>. Acessado em 23 de Novembro de 2019.

WONDERLIC TEST SAMPLE. **The Wonderlic Test and the NFL Combine.** Disponível em: <<https://wonderlictestsample.com/the-wonderlic-test-and-the-nfl-combine/>>. Acessado em 15 de Novembro de 2019.

APÊNDICE A – Rating com 16 variáveis independentes com multicolinearidade

Regressão

Estatística Descritiva

	Média	Erro Desvio	N
NFL_Rate	72,57	16,57	220,00
Comb_Height (in)	74,89	1,63	222,00
Comb_Weight (lbs)	219,91	14,04	222,00
Com_Wonderlic	27,25	6,78	122,00
Comb_40 Yard	4,85	,19	200,00
Comb_Vert Leap (in)	31,27	3,29	176,00
Comb_Broad Jump (in)	109,52	6,91	173,00
Comb_Shuttle	4,31	,17	169,00
Comb_3Cone	7,17	,27	114,00
College_G	38,61	9,26	202,00
College_Comp	594,40	266,73	202,00
College_Att	989,69	409,15	202,00
College_Yds	7628,64	3185,49	202,00
College_TD	55,86	27,51	202,00
College_INT	29,31	12,62	202,00
College_RAT	134,49	29,00	202,00

Correlações

		NFL_Rate	Comb_Height (in)	Comb_Weight (lbs)	Com_Wonderlic	Comb_40 Yard	Comb_Vert Leap (in)	Comb_Broad Jump (in)
Correlação de Pearson	NFL_Rate	1,00	,05	,05	,22	-,02	,02	,14
	Comb_Height (in)	,05	1,00	,62	,17	,25	-,15	-,05
	Comb_Weight (lbs)	,05	,62	1,00	,09	,01	,10	,11
	Com_Wonderlic	,22	,17	,09	1,00	,20	-,19	-,15
	Comb_40 Yard	-,02	,25	,01	,20	1,00	-,73	-,67
	Comb_Vert Leap (in)	,02	-,15	,10	-,19	-,73	1,00	,75
	Comb_Broad Jump (in)	,14	-,05	,11	-,15	-,67	,75	1,00
	Comb_Shuttle	-,03	,17	,00	-,12	,46	-,48	-,45
	Comb_3Cone	-,11	,05	-,10	-,25	,44	-,35	-,47
	College_G	,16	-,07	-,07	,11	,01	-,05	-,06
	College_Comp	,24	-,04	,02	,13	,09	,01	-,03
	College_Att	,20	-,06	,00	,11	,10	-,01	-,04
	College_Yds	,21	-,07	,00	,09	,05	,03	-,02
	College_TD	,21	-,09	,02	,07	,00	,06	-,01
	College_INT	-,02	-,09	-,19	,02	,20	-,11	-,17
	College_RAT	,18	-,04	,02	,04	-,11	,03	,09
Sig. (1 extremidade)	NFL_Rate	.	,21	,21	,01	,37	,42	,03
	Comb_Height (in)	,21	.	,00	,03	,00	,02	,25
	Comb_Weight (lbs)	,21	,00	.	,17	,42	,09	,07
	Com_Wonderlic	,01	,03	,17	.	,01	,03	,07
	Comb_40 Yard	,37	,00	,42	,01	.	,00	,00
	Comb_Vert Leap (in)	,42	,02	,09	,03	,00	.	,00
	Comb_Broad Jump (in)	,03	,25	,07	,07	,00	,00	.
	Comb_Shuttle	,34	,01	,49	,14	,00	,00	,00
	Comb_3Cone	,14	,31	,15	,01	,00	,00	,00
	College_G	,01	,17	,16	,13	,44	,28	,22
	College_Comp	,00	,27	,38	,09	,11	,47	,35
	College_Att	,00	,21	,48	,12	,09	,47	,30
	College_Yds	,00	,15	,49	,18	,24	,34	,41
	College_TD	,00	,10	,36	,22	,48	,22	,44

	College_INT	,40	,09	,00	,42	,00	,09	,02
	College_RAT	,01	,28	,39	,33	,07	,33	,14
N	NFL_Rate	220,00	220,00	220,00	121,00	198,00	174,00	171,00
	Comb_Height (in)	220,00	222,00	222,00	122,00	200,00	176,00	173,00
	Comb_Weight (lbs)	220,00	222,00	222,00	122,00	200,00	176,00	173,00
	Com_Wonderlic	121,00	122,00	122,00	122,00	121,00	98,00	97,00
	Comb_40 Yard	198,00	200,00	200,00	121,00	200,00	172,00	168,00
	Comb_Vert Leap (in)	174,00	176,00	176,00	98,00	172,00	176,00	169,00
	Comb_Broad Jump (in)	171,00	173,00	173,00	97,00	168,00	169,00	173,00
	Comb_Shuttle	167,00	169,00	169,00	93,00	165,00	164,00	161,00
	Comb_3Cone	112,00	114,00	114,00	92,00	112,00	109,00	108,00
	College_G	200,00	202,00	202,00	114,00	182,00	158,00	154,00
	College_Comp	200,00	202,00	202,00	114,00	182,00	158,00	154,00
	College_Att	200,00	202,00	202,00	114,00	182,00	158,00	154,00
	College_Yds	200,00	202,00	202,00	114,00	182,00	158,00	154,00
	College_TD	200,00	202,00	202,00	114,00	182,00	158,00	154,00
	College_INT	200,00	202,00	202,00	114,00	182,00	158,00	154,00
	College_RAT	200,00	202,00	202,00	114,00	182,00	158,00	154,00

Correlações

		Comb_Shuttle	Comb_3Cone	College_G	College_Comp	College_Att	College_Yds	College_TD
Correlação de Pearson	NFL_Rate	-,03	-,11	,16	,24	,20	,21	,21
	Comb_Height (in)	,17	,05	-,07	-,04	-,06	-,07	-,09
	Comb_Weight (lbs)	,00	-,10	-,07	,02	,00	,00	,02
	Com_Wonderlic	-,12	-,25	,11	,13	,11	,09	,07
	Comb_40 Yard	,46	,44	,01	,09	,10	,05	,00
	Comb_Vert Leap (in)	-,48	-,35	-,05	,01	-,01	,03	,06
	Comb_Broad Jump (in)	-,45	-,47	-,06	-,03	-,04	-,02	-,01
	Comb_Shuttle	1,00	,61	-,01	,00	,01	-,02	-,05
	Comb_3Cone	,61	1,00	-,06	-,04	,00	-,02	-,01
	College_G	-,01	-,06	1,00	,62	,65	,64	,56
	College_Comp	,00	-,04	,62	1,00	,98	,98	,90
	College_Att	,01	,00	,65	,98	1,00	,97	,87

	College_Yds	-,02	-,02	,64	,98	,97	1,00	,93
	College_TD	-,05	-,01	,56	,90	,87	,93	1,00
	College_INT	,12	,13	,51	,67	,76	,68	,52
	College_RAT	-,19	-,22	,34	,31	,25	,35	,42
Sig. (1 extremidade)	NFL_Rate	,34	,14	,01	,00	,00	,00	,00
	Comb_Height (in)	,01	,31	,17	,27	,21	,15	,10
	Comb_Weight (lbs)	,49	,15	,16	,38	,48	,49	,36
	Com_Wonderlic	,14	,01	,13	,09	,12	,18	,22
	Comb_40 Yard	,00	,00	,44	,11	,09	,24	,48
	Comb_Vert Leap (in)	,00	,00	,28	,47	,47	,34	,22
	Comb_Broad Jump (in)	,00	,00	,22	,35	,30	,41	,44
	Comb_Shuttle	.	,00	,47	,49	,43	,41	,26
	Comb_3Cone	,00	.	,27	,34	,48	,40	,47
	College_G	,47	,27	.	,00	,00	,00	,00
	College_Comp	,49	,34	,00	.	,00	,00	,00
	College_Att	,43	,48	,00	,00	.	,00	,00
	College_Yds	,41	,40	,00	,00	,00	.	,00
	College_TD	,26	,47	,00	,00	,00	,00	.
	College_INT	,08	,09	,00	,00	,00	,00	,00
	College_RAT	,01	,01	,00	,00	,00	,00	,00
N	NFL_Rate	167,00	112,00	200,00	200,00	200,00	200,00	200,00
	Comb_Height (in)	169,00	114,00	202,00	202,00	202,00	202,00	202,00
	Comb_Weight (lbs)	169,00	114,00	202,00	202,00	202,00	202,00	202,00
	Com_Wonderlic	93,00	92,00	114,00	114,00	114,00	114,00	114,00
	Comb_40 Yard	165,00	112,00	182,00	182,00	182,00	182,00	182,00
	Comb_Vert Leap (in)	164,00	109,00	158,00	158,00	158,00	158,00	158,00
	Comb_Broad Jump (in)	161,00	108,00	154,00	154,00	154,00	154,00	154,00
	Comb_Shuttle	169,00	112,00	151,00	151,00	151,00	151,00	151,00
	Comb_3Cone	112,00	114,00	106,00	106,00	106,00	106,00	106,00
	College_G	151,00	106,00	202,00	202,00	202,00	202,00	202,00
	College_Comp	151,00	106,00	202,00	202,00	202,00	202,00	202,00
	College_Att	151,00	106,00	202,00	202,00	202,00	202,00	202,00
	College_Yds	151,00	106,00	202,00	202,00	202,00	202,00	202,00

College_TD	151,00	106,00	202,00	202,00	202,00	202,00	202,00
College_INT	151,00	106,00	202,00	202,00	202,00	202,00	202,00
College_RAT	151,00	106,00	202,00	202,00	202,00	202,00	202,00

Correlações

		College_INT	College_RAT
Correlação de Pearson	NFL_Rate	-,02	,18
	Comb_Height (in)	-,09	-,04
	Comb_Weight (lbs)	-,19	,02
	Com_Wonderlic	,02	,04
	Comb_40 Yard	,20	-,11
	Comb_Vert Leap (in)	-,11	,03
	Comb_Broad Jump (in)	-,17	,09
	Comb_Shuttle	,12	-,19
	Comb_3Cone	,13	-,22
	College_G	,51	,34
	College_Comp	,67	,31
	College_Att	,76	,25
	College_Yds	,68	,35
	College_TD	,52	,42
	College_INT	1,00	,04
	College_RAT	,04	1,00
	Sig. (1 extremidade)	NFL_Rate	,40
Comb_Height (in)		,09	,28
Comb_Weight (lbs)		,00	,39
Com_Wonderlic		,42	,33
Comb_40 Yard		,00	,07
Comb_Vert Leap (in)		,09	,33
Comb_Broad Jump (in)		,02	,14
Comb_Shuttle		,08	,01
Comb_3Cone		,09	,01
College_G		,00	,00
College_Comp		,00	,00
College_Att		,00	,00
College_Yds		,00	,00
College_TD		,00	,00
College_INT		.	,29
College_RAT		,29	.
N		NFL_Rate	200,00
	Comb_Height (in)	202,00	202,00
	Comb_Weight (lbs)	202,00	202,00

Com_Wonderlic	114,00	114,00
Comb_40 Yard	182,00	182,00
Comb_Vert Leap (in)	158,00	158,00
Comb_Broad Jump (in)	154,00	154,00
Comb_Shuttle	151,00	151,00
Comb_3Cone	106,00	106,00
College_G	202,00	202,00
College_Comp	202,00	202,00
College_Att	202,00	202,00
College_Yds	202,00	202,00
College_TD	202,00	202,00
College_INT	202,00	202,00
College_RAT	202,00	202,00

Variáveis Inseridas/Removidas^a

Modelo	Variáveis inseridas	Variáveis removidas	Método
1	College_RAT, Comb_Weight (lbs), Com_Wonderlic, College_INT, Comb_Vert Leap (in), Comb_3Cone, College_G, Comb_Height (in), Comb_Shuttle, College_TD, Comb_40 Yard, Comb_Broad Jump (in), College_Comp, College_Yds, College_Att ^b		Inserir

a. Variável Dependente: NFL_Rate

b. Todas as variáveis solicitadas inseridas.

Resumo do modelo^b

Modelo	R	R quadrado	R quadrado ajustado	Erro padrão da estimativa	Estatísticas de mudança	
					Mudança de R quadrado	Mudança F
1	,44 ^a	,19	,03	16,29	,19	1,21

Resumo do modelo^b

Modelo	Estatísticas de mudança		
	df1	df2	Sig. Mudança F
1	15,00	76,00	,28

a. Preditores: (Constante), College_RAT, Comb_Weight (lbs), Com_Wonderlic, College_INT, Comb_Vert Leap (in), Comb_3Cone, College_G, Comb_Height (in), Comb_Shuttle, College_TD, Comb_40 Yard, Comb_Broad Jump (in), College_Comp, College_Yds, College_Att

b. Variável Dependente: NFL_Rate

ANOVA^a

Modelo		Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
1	Regressão	4834,06	15,00	322,27	1,21	,28 ^b
	Resíduo	20159,64	76,00	265,26		
	Total	24993,70	91,00			

a. Variável Dependente: NFL_Rate

b. Preditores: (Constante), College_RAT, Comb_Weight (lbs), Com_Wonderlic, College_INT, Comb_Vert Leap (in), Comb_3Cone, College_G, Comb_Height (in), Comb_Shuttle, College_TD, Comb_40 Yard, Comb_Broad Jump (in), College_Comp, College_Yds, College_Att

Coefficientes^a

Modelo		Coefficients não padronizados		Coefficients padronizados		Sig.
		B	Erro Erro	Beta	t	
1	(Constante)	-66,80	135,40		-,49	,62
	Comb_Height (in)	,22	1,46	,02	,15	,88
	Comb_Weight (lbs)	-,07	,17	-,06	-,39	,70
	Com_Wonderlic	,49	,29	,20	1,67	,10
	Comb_40 Yard	-,50	15,56	-,01	-,03	,97

Comb_Vert Leap (in)	-,68	,96	-,13	-,71	,48
Comb_Broad Jump (in)	,72	,42	,30	1,70	,09
Comb_Shuttle	5,24	13,64	,06	,38	,70
Comb_3Cone	4,78	9,50	,08	,50	,62
College_G	,08	,26	,05	,31	,76
College_Comp	,04	,05	,57	,73	,47
College_Att	,01	,03	,23	,27	,78
College_Yds	,00	,00	-,38	-,54	,59
College_TD	-,03	,20	-,05	-,13	,89
College_INT	-,40	,28	-,31	-1,46	,15
College_RAT	,05	,07	,10	,73	,47

Coefficientes^a

Modelo	Ordem zero	Correlações			Estatísticas de colinearidade	
		Parcial	Parte	Tolerância	VIF	
1 (Constante)						
Comb_Height (in)	,05	,02	,02	,52	1,94	
Comb_Weight (lbs)	,05	-,04	-,04	,52	1,93	
Com_Wonderlic	,22	,19	,17	,75	1,33	
Comb_40 Yard	-,02	,00	,00	,35	2,88	
Comb_Vert Leap (in)	,02	-,08	-,07	,29	3,40	
Comb_Broad Jump (in)	,14	,19	,17	,34	2,94	
Comb_Shuttle	-,03	,04	,04	,51	1,96	
Comb_3Cone	-,11	,06	,05	,44	2,25	
College_G	,16	,04	,03	,50	1,99	
College_Comp	,24	,08	,07	,02	57,01	
College_Att	,20	,03	,03	,01	66,80	
College_Yds	,21	-,06	-,06	,02	47,44	
College_TD	,21	-,02	-,01	,09	10,67	
College_INT	-,02	-,17	-,15	,24	4,20	
College_RAT	,18	,08	,08	,63	1,59	

a. Variável Dependente: NFL_Rate

Diagnóstico entre casos^a

Número do caso	Erro Resíduo	NFL_Rate	Valor previsto	Resíduo
115	-3,20	8,80	60,95	-52,15
117	-3,47	,60	57,04	-56,44

126	-3,19	16,40	68,42	-52,02
206	4,02	152,10	86,67	65,43

a. Variável Dependente: NFL_Rate

Estadísticas de resíduos^a

	Mínimo	Máximo	Média	Erro Desvio	N
Valor previsto	48,93	88,24	72,35	7,85	76,00
Erro Valor previsto	-3,24	2,15	-,03	1,08	76,00
Erro padrão do valor previsto	4,25	22,52	6,74	2,21	76,00
Valor previsto ajustado	51,41	94,00	72,68	8,68	74,00
Resíduo	-56,44	65,43	-,33	19,01	75,00
Erro Resíduo	-3,47	4,02	-,02	1,17	75,00
Resíduos Resíduo de Estud.	-3,66	4,39	-,01	1,29	74,00
Resíduos de Estud.	-64,92	78,15	-,18	23,01	74,00
Resíduos de Estud.	-4,01	5,05	-,02	1,36	74,00
Mahal. Distância	5,22	173,07	16,26	19,11	76,00
Distância de Cook	,00	,23	,02	,04	74,00
Valor de ponto alavanca centralizado	,06	1,90	,18	,21	76,00

a. Variável Dependente: NFL_Rate

GRÁFICOS

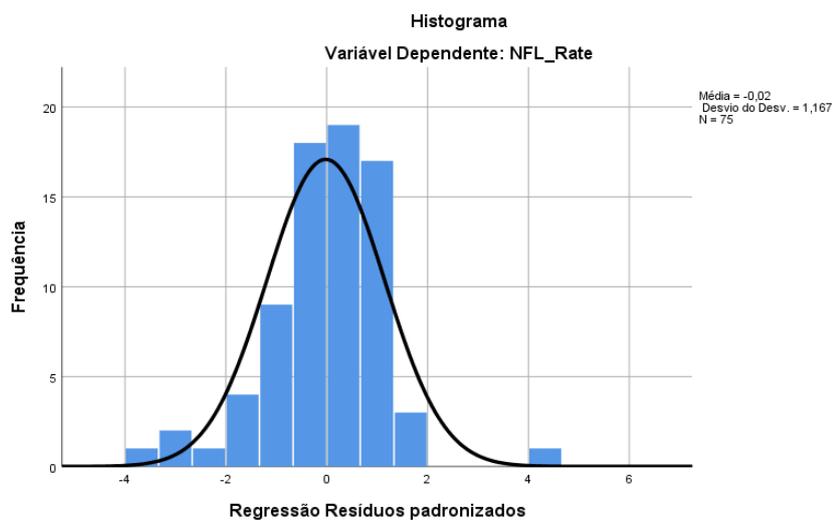
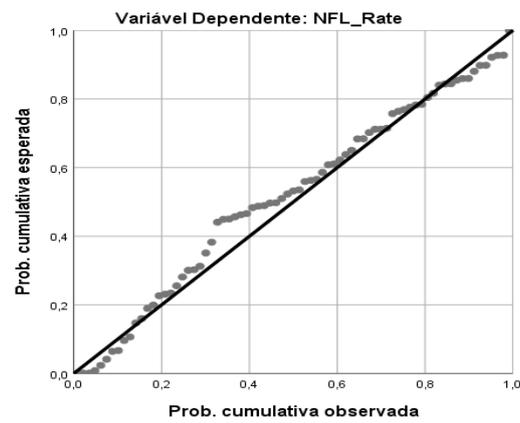
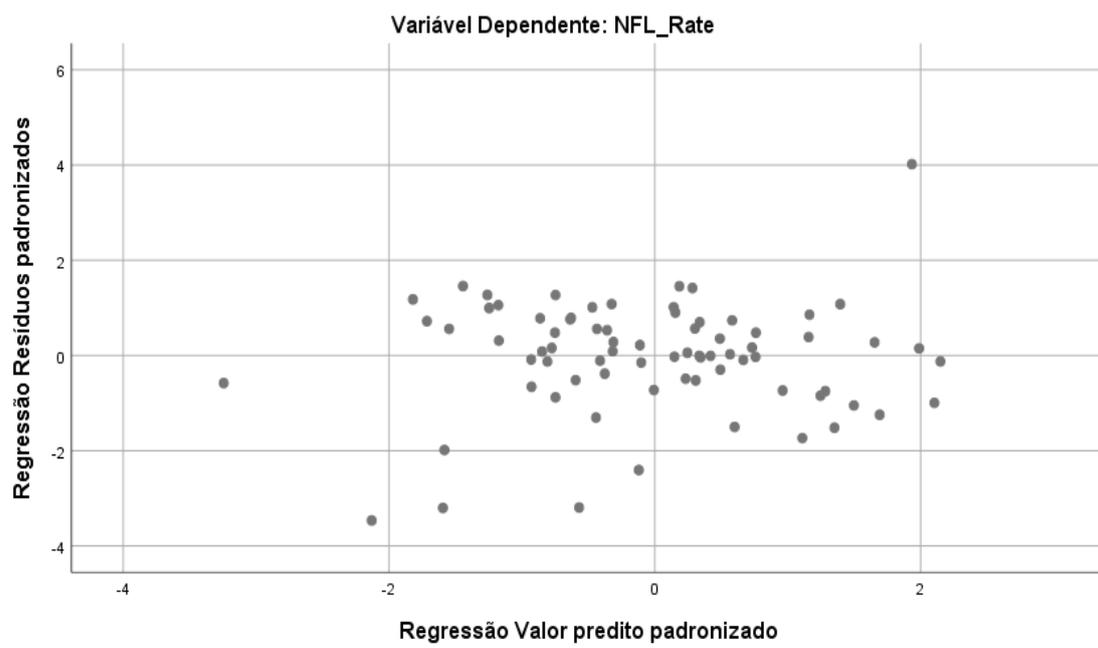


Gráfico P-P Normal de Regressão Resíduos padronizados**Gráfico de dispersão**

APÊNDICE B – Rating com 12 variáveis independentes

Regressão

Estatística Descritiva

	Média	Erro Desvio	N
NFL_Rate	72,57	16,57	220,00
Comb_Height (in)	74,89	1,63	222,00
Comb_Weight (lbs)	219,91	14,04	222,00
Com_Wonderlic	27,25	6,78	122,00
Comb_40 Yard	4,85	,19	200,00
Comb_Vert Leap (in)	31,27	3,29	176,00
Comb_Broad Jump (in)	109,52	6,91	173,00
Comb_Shuttle	4,31	,17	169,00
Comb_3Cone	7,17	,27	114,00
College_G	38,61	9,26	202,00
College_TD	55,86	27,51	202,00
College_INT	29,31	12,62	202,00
College_RAT	134,49	29,00	202,00

Correlações

		NFL_Rate	Comb_Height (in)	Comb_Weight (lbs)	Com_Wonderlic	Comb_40 Yard
Correlação de Pearson	NFL_Rate	1,00	,05	,05	,22	-,02
	Comb_Height (in)	,05	1,00	,62	,17	,25
	Comb_Weight (lbs)	,05	,62	1,00	,09	,01
	Com_Wonderlic	,22	,17	,09	1,00	,20
	Comb_40 Yard	-,02	,25	,01	,20	1,00
	Comb_Vert Leap (in)	,02	-,15	,10	-,19	-,73
	Comb_Broad Jump (in)	,14	-,05	,11	-,15	-,67
	Comb_Shuttle	-,03	,17	,00	-,12	,46
	Comb_3Cone	-,11	,05	-,10	-,25	,44
	College_G	,16	-,07	-,07	,11	,01
	College_TD	,21	-,09	,02	,07	,00
	College_INT	-,02	-,09	-,19	,02	,20
	College_RAT	,18	-,04	,02	,04	-,11
	Sig. (1 extremidade)	NFL_Rate	.	,21	,21	,01
Comb_Height (in)		,21	.	,00	,03	,00
Comb_Weight (lbs)		,21	,00	.	,17	,42

	Com_Wonderlic	,01	,03	,17	.	,01
	Comb_40 Yard	,37	,00	,42	,01	.
	Comb_Vert Leap (in)	,42	,02	,09	,03	,00
	Comb_Broad Jump (in)	,03	,25	,07	,07	,00
	Comb_Shuttle	,34	,01	,49	,14	,00
	Comb_3Cone	,14	,31	,15	,01	,00
	College_G	,01	,17	,16	,13	,44
	College_TD	,00	,10	,36	,22	,48
	College_INT	,40	,09	,00	,42	,00
	College_RAT	,01	,28	,39	,33	,07
N	NFL_Rate	220,00	220,00	220,00	121,00	198,00
	Comb_Height (in)	220,00	222,00	222,00	122,00	200,00
	Comb_Weight (lbs)	220,00	222,00	222,00	122,00	200,00
	Com_Wonderlic	121,00	122,00	122,00	122,00	121,00
	Comb_40 Yard	198,00	200,00	200,00	121,00	200,00
	Comb_Vert Leap (in)	174,00	176,00	176,00	98,00	172,00
	Comb_Broad Jump (in)	171,00	173,00	173,00	97,00	168,00
	Comb_Shuttle	167,00	169,00	169,00	93,00	165,00
	Comb_3Cone	112,00	114,00	114,00	92,00	112,00
	College_G	200,00	202,00	202,00	114,00	182,00
	College_TD	200,00	202,00	202,00	114,00	182,00
	College_INT	200,00	202,00	202,00	114,00	182,00
	College_RAT	200,00	202,00	202,00	114,00	182,00

Correlações

		Comb_Vert Leap (in)	Comb_Broad Jump (in)	Comb_Shuttle	Comb_3Cone	College_G
Correlação de Pearson	NFL_Rate	,02	,14	-,03	-,11	,16
	Comb_Height (in)	-,15	-,05	,17	,05	-,07
	Comb_Weight (lbs)	,10	,11	,00	-,10	-,07
	Com_Wonderlic	-,19	-,15	-,12	-,25	,11
	Comb_40 Yard	-,73	-,67	,46	,44	,01
	Comb_Vert Leap (in)	1,00	,75	-,48	-,35	-,05
	Comb_Broad Jump (in)	,75	1,00	-,45	-,47	-,06
	Comb_Shuttle	-,48	-,45	1,00	,61	-,01
	Comb_3Cone	-,35	-,47	,61	1,00	-,06
	College_G	-,05	-,06	-,01	-,06	1,00
	College_TD	,06	-,01	-,05	-,01	,56
	College_INT	-,11	-,17	,12	,13	,51
	College_RAT	,03	,09	-,19	-,22	,34
Sig. (1 extremidade)	NFL_Rate	,42	,03	,34	,14	,01
	Comb_Height (in)	,02	,25	,01	,31	,17

	Comb_Weight (lbs)	,09	,07	,49	,15	,16
	Com_Wonderlic	,03	,07	,14	,01	,13
	Comb_40 Yard	,00	,00	,00	,00	,44
	Comb_Vert Leap (in)	.	,00	,00	,00	,28
	Comb_Broad Jump (in)	,00	.	,00	,00	,22
	Comb_Shuttle	,00	,00	.	,00	,47
	Comb_3Cone	,00	,00	,00	.	,27
	College_G	,28	,22	,47	,27	.
	College_TD	,22	,44	,26	,47	,00
	College_INT	,09	,02	,08	,09	,00
	College_RAT	,33	,14	,01	,01	,00
N	NFL_Rate	174,00	171,00	167,00	112,00	200,00
	Comb_Height (in)	176,00	173,00	169,00	114,00	202,00
	Comb_Weight (lbs)	176,00	173,00	169,00	114,00	202,00
	Com_Wonderlic	98,00	97,00	93,00	92,00	114,00
	Comb_40 Yard	172,00	168,00	165,00	112,00	182,00
	Comb_Vert Leap (in)	176,00	169,00	164,00	109,00	158,00
	Comb_Broad Jump (in)	169,00	173,00	161,00	108,00	154,00
	Comb_Shuttle	164,00	161,00	169,00	112,00	151,00
	Comb_3Cone	109,00	108,00	112,00	114,00	106,00
	College_G	158,00	154,00	151,00	106,00	202,00
	College_TD	158,00	154,00	151,00	106,00	202,00
	College_INT	158,00	154,00	151,00	106,00	202,00
	College_RAT	158,00	154,00	151,00	106,00	202,00

Correlações

		College_TD	College_INT	College_RAT
Correlação de Pearson	NFL_Rate	,21	-,02	,18
	Comb_Height (in)	-,09	-,09	-,04
	Comb_Weight (lbs)	,02	-,19	,02
	Com_Wonderlic	,07	,02	,04
	Comb_40 Yard	,00	,20	-,11
	Comb_Vert Leap (in)	,06	-,11	,03
	Comb_Broad Jump (in)	-,01	-,17	,09
	Comb_Shuttle	-,05	,12	-,19
	Comb_3Cone	-,01	,13	-,22
	College_G	,56	,51	,34
	College_TD	1,00	,52	,42
	College_INT	,52	1,00	,04
	College_RAT	,42	,04	1,00
Sig. (1 extremidade)	NFL_Rate	,00	,40	,01

	Comb_Height (in)	,10	,09	,28
	Comb_Weight (lbs)	,36	,00	,39
	Com_Wonderlic	,22	,42	,33
	Comb_40 Yard	,48	,00	,07
	Comb_Vert Leap (in)	,22	,09	,33
	Comb_Broad Jump (in)	,44	,02	,14
	Comb_Shuttle	,26	,08	,01
	Comb_3Cone	,47	,09	,01
	College_G	,00	,00	,00
	College_TD	.	,00	,00
	College_INT	,00	.	,29
	College_RAT	,00	,29	.
N	NFL_Rate	200,00	200,00	200,00
	Comb_Height (in)	202,00	202,00	202,00
	Comb_Weight (lbs)	202,00	202,00	202,00
	Com_Wonderlic	114,00	114,00	114,00
	Comb_40 Yard	182,00	182,00	182,00
	Comb_Vert Leap (in)	158,00	158,00	158,00
	Comb_Broad Jump (in)	154,00	154,00	154,00
	Comb_Shuttle	151,00	151,00	151,00
	Comb_3Cone	106,00	106,00	106,00
	College_G	202,00	202,00	202,00
	College_TD	202,00	202,00	202,00
	College_INT	202,00	202,00	202,00
	College_RAT	202,00	202,00	202,00

Variáveis Inseridas/Removidas^a

Modelo	Variáveis inseridas	Variáveis removidas	Método
1	College_RAT, Comb_Weight (lbs), Com_Wonderlic, College_INT, Comb_Vert Leap (in), Comb_3Cone, College_G, Comb_Height (in), Comb_Shuttle, College_TD, Comb_40 Yard, Comb_Broad Jump (in) ^b		Inserir

a. Variável Dependente: NFL_Rate

b. Todas as variáveis solicitadas inseridas.

Resumo do modelo^b

Modelo	R	R quadrado	R quadrado ajustado	Erro padrão da estimativa	Estatísticas de mudança	
					Mudança de R quadrado	Mudança F
1	,40 ^a	,16	,03	16,29	,16	1,26

Resumo do modelo^b

Modelo	Estatísticas de mudança			Sig. Mudança F
	df1	df2	Sig.	
1	12,00	79,00	,26	1,69

a. Preditores: (Constante), College_RAT, Comb_Weight (lbs), Com_Wonderlic, College_INT, Comb_Vert Leap (in), Comb_3Cone, College_G, Comb_Height (in), Comb_Shuttle, College_TD, Comb_40 Yard, Comb_Broad Jump (in)

b. Variável Dependente: NFL_Rate

ANOVA^a

Modelo		Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
1	Regressão	4019,93	12,00	334,99	1,26	,26 ^b
	Resíduo	20973,77	79,00	265,49		
	Total	24993,70	91,00			

a. Variável Dependente: NFL_Rate

b. Preditores: (Constante), College_RAT, Comb_Weight (lbs), Com_Wonderlic, College_INT, Comb_Vert Leap (in), Comb_3Cone, College_G, Comb_Height (in), Comb_Shuttle, College_TD, Comb_40 Yard, Comb_Broad Jump (in)

Coeficientes^a

Modelo		Coeficientes não padronizados		Coeficientes padronizados	t	Sig.
		B	Erro Erro	Beta		
1	(Constante)	-90,77	134,39		-,68	,50
	Comb_Height (in)	,17	1,46	,02	,12	,91
	Comb_Weight (lbs)	-,04	,17	-,03	-,21	,83
	Com_Wonderlic	,54	,29	,22	1,87	,07
	Comb_40 Yard	4,27	15,16	,05	,28	,78
	Comb_Vert Leap (in)	-,71	,95	-,14	-,75	,45
	Comb_Broad Jump (in)	,82	,42	,34	1,96	,05
	Comb_Shuttle	7,58	13,51	,08	,56	,58
	Comb_3Cone	1,68	9,17	,03	,18	,85
	College_G	,16	,24	,09	,66	,51
	College_TD	,14	,09	,23	1,56	,12
	College_INT	-,23	,18	-,18	-1,25	,21
	College_RAT	,03	,07	,05	,40	,69

Coeficientes^a

Modelo		Correlações			Estatísticas de colinearidade	
		Ordem zero	Parcial	Parte	Tolerância	VIF
1	(Constante)					
	Comb_Height (in)	,05	,01	,01	,52	1,93
	Comb_Weight (lbs)	,05	-,02	-,02	,53	1,87
	Com_Wonderlic	,22	,21	,19	,76	1,31
	Comb_40 Yard	-,02	,03	,03	,37	2,73
	Comb_Vert Leap (in)	,02	-,08	-,08	,30	3,33
	Comb_Broad Jump (in)	,14	,22	,20	,35	2,84

Comb_Shuttle	-,03	,06	,06	,52	1,91
Comb_3Cone	-,11	,02	,02	,48	2,09
College_G	,16	,07	,07	,57	1,75
College_TD	,21	,17	,16	,48	2,07
College_INT	-,02	-,14	-,13	,54	1,86
College_RAT	,18	,05	,04	,69	1,44

a. Variável Dependente: NFL_Rate

Diagnóstico entre casos^a

Número do caso	Erro Resíduo	NFL_Rate	Valor previsto	Resíduo
115	-3,33	8,80	63,12	-54,32
117	-3,57	,60	58,71	-58,11
206	4,16	152,10	84,36	67,74

a. Variável Dependente: NFL_Rate

Estatísticas de resíduos^a

	Mínimo	Máximo	Média	Erro Desvio	N
Valor previsto	54,72	90,11	72,23	7,00	76,00
Erro Valor previsto	-2,69	2,64	-,05	1,05	76,00
Erro padrão do valor previsto	3,96	21,73	6,06	2,20	76,00
Valor previsto ajustado	52,89	90,21	72,34	7,53	74,00
Resíduo	-58,11	67,74	-,21	19,15	75,00
Erro Resíduo	-3,57	4,16	-,01	1,18	75,00
Resíduos Resíduo	-3,76	4,48	,00	1,27	74,00
de Estud.	-66,72	78,58	,17	22,29	74,00
Resíduos de Estud.	-4,12	5,15	,00	1,34	74,00
Mahal. Distância	4,39	160,90	13,23	17,96	76,00
Distância de Cook	,00	,25	,02	,04	74,00
Valor de ponto alavanca centralizado	,05	1,77	,15	,20	76,00

a. Variável Dependente: NFL_Rate

GRÁFICOS

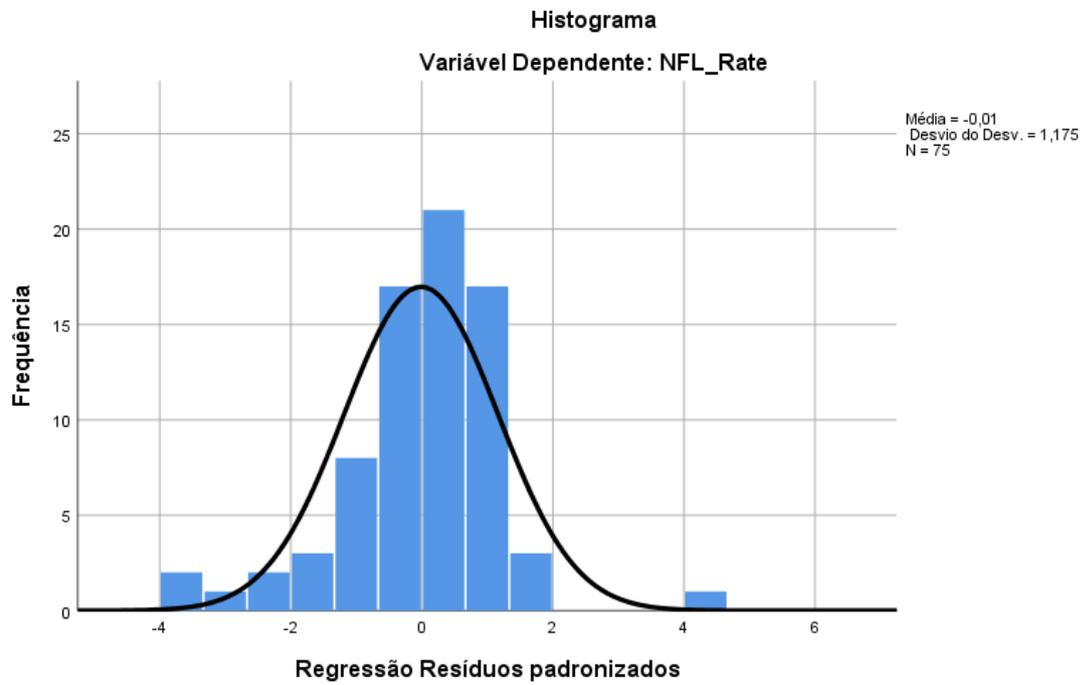
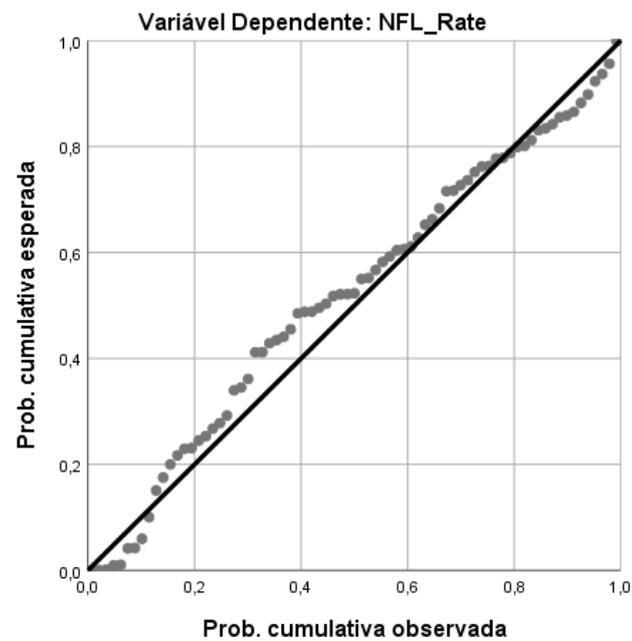
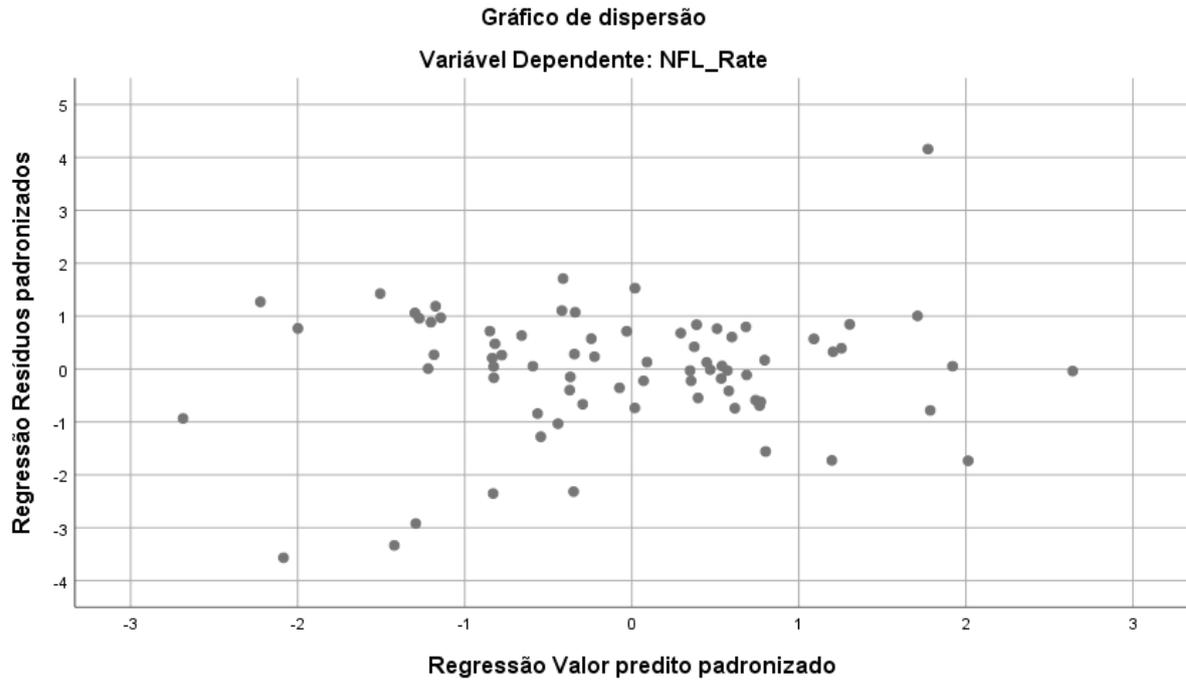


Gráfico P-P Normal de Regressão Resíduos padronizados





APÊNDICE C – Draft

Estatística Descritiva

	Média	Erro Desvio	N
Draft_Selected	87,14	74,14	196,00
Comb_Height (in)	74,89	1,63	222,00
Comb_Weight (lbs)	219,91	14,04	222,00
Comb_Wonderlic	27,25	6,78	122,00
Comb_40 Yard	4,85	,19	200,00
Comb_Vert Leap (in)	31,27	3,29	176,00
Comb_Broad Jump (in)	109,52	6,91	173,00
Comb_Shuttle	4,31	,17	169,00
Comb_3Cone	7,17	,27	114,00
College_G	38,61	9,26	202,00
College_TD	55,86	27,51	202,00
College_INT	29,31	12,62	202,00
College_RAT	134,49	29,00	202,00

Correlações

		Draft_Selecte d	Comb_Heigh t (in)	Comb_Weig ht (lbs)	Comb_Wond erlic	Comb_40 Ya rd
Correlação de Pearson	Draft_Selected	1,00	-,21	-,27	-,06	,22
	Comb_Height (in)	-,21	1,00	,62	,17	,25
	Comb_Weight (lbs)	-,27	,62	1,00	,09	,01
	Comb_Wonderlic	-,06	,17	,09	1,00	,20
	Comb_40 Yard	,22	,25	,01	,20	1,00
	Comb_Vert Leap (in)	-,28	-,15	,10	-,19	-,73
	Comb_Broad Jump (in)	-,30	-,05	,11	-,15	-,67
	Comb_Shuttle	,16	,17	,00	-,12	,46
	Comb_3Cone	,33	,05	-,10	-,25	,44
	College_G	,10	-,07	-,07	,11	,01
	College_TD	-,16	-,09	,02	,07	,00
	College_INT	,18	-,09	-,19	,02	,20
	College_RAT	-,22	-,04	,02	,04	-,11
Sig. (1 extremidade)	Draft_Selected	.	,00	,00	,26	,00
	Comb_Height (in)	,00	.	,00	,03	,00
	Comb_Weight (lbs)	,00	,00	.	,17	,42
	Comb_Wonderlic	,26	,03	,17	.	,01

	Comb_40 Yard	,00	,00	,42	,01	.
	Comb_Vert Leap (in)	,00	,02	,09	,03	,00
	Comb_Broad Jump (in)	,00	,25	,07	,07	,00
	Comb_Shuttle	,03	,01	,49	,14	,00
	Comb_3Cone	,00	,31	,15	,01	,00
	College_G	,10	,17	,16	,13	,44
	College_TD	,02	,10	,36	,22	,48
	College_INT	,01	,09	,00	,42	,00
	College_RAT	,00	,28	,39	,33	,07
N	Draft_Selected	196,00	196,00	196,00	113,00	175,00
	Comb_Height (in)	196,00	222,00	222,00	122,00	200,00
	Comb_Weight (lbs)	196,00	222,00	222,00	122,00	200,00
	Comb_Wonderlic	113,00	122,00	122,00	122,00	121,00
	Comb_40 Yard	175,00	200,00	200,00	121,00	200,00
	Comb_Vert Leap (in)	151,00	176,00	176,00	98,00	172,00
	Comb_Broad Jump (in)	149,00	173,00	173,00	97,00	168,00
	Comb_Shuttle	147,00	169,00	169,00	93,00	165,00
	Comb_3Cone	101,00	114,00	114,00	92,00	112,00
	College_G	182,00	202,00	202,00	114,00	182,00
	College_TD	182,00	202,00	202,00	114,00	182,00
	College_INT	182,00	202,00	202,00	114,00	182,00
	College_RAT	182,00	202,00	202,00	114,00	182,00

Correlações

		Comb_Vert Leap (in)	Comb_Broad Jump (in)	Comb_Shuttl e	Comb_3Con e	College_G
Correlação de Pearson	Draft_Selected	-,28	-,30	,16	,33	,10
	Comb_Height (in)	-,15	-,05	,17	,05	-,07
	Comb_Weight (lbs)	,10	,11	,00	-,10	-,07
	Comb_Wonderlic	-,19	-,15	-,12	-,25	,11
	Comb_40 Yard	-,73	-,67	,46	,44	,01
	Comb_Vert Leap (in)	1,00	,75	-,48	-,35	-,05
	Comb_Broad Jump (in)	,75	1,00	-,45	-,47	-,06
	Comb_Shuttle	-,48	-,45	1,00	,61	-,01
	Comb_3Cone	-,35	-,47	,61	1,00	-,06
	College_G	-,05	-,06	-,01	-,06	1,00
	College_TD	,06	-,01	-,05	-,01	,56
	College_INT	-,11	-,17	,12	,13	,51

	College_RAT	,03	,09	-,19	-,22	,34	
Sig. (1 extremidade)	Draft_Selected	,00	,00	,03	,00	,10	
	Comb_Height (in)	,02	,25	,01	,31	,17	
	Comb_Weight (lbs)	,09	,07	,49	,15	,16	
	Comb_Wonderlic	,03	,07	,14	,01	,13	
	Comb_40 Yard	,00	,00	,00	,00	,44	
	Comb_Vert Leap (in)	.	,00	,00	,00	,28	
	Comb_Broad Jump (in)	,00	.	,00	,00	,22	
	Comb_Shuttle	,00	,00	.	,00	,47	
	Comb_3Cone	,00	,00	,00	.	,27	
	College_G	,28	,22	,47	,27	.	
	College_TD	,22	,44	,26	,47	,00	
	College_INT	,09	,02	,08	,09	,00	
	College_RAT	,33	,14	,01	,01	,00	
	N	Draft_Selected	151,00	149,00	147,00	101,00	182,00
		Comb_Height (in)	176,00	173,00	169,00	114,00	202,00
Comb_Weight (lbs)		176,00	173,00	169,00	114,00	202,00	
Comb_Wonderlic		98,00	97,00	93,00	92,00	114,00	
Comb_40 Yard		172,00	168,00	165,00	112,00	182,00	
Comb_Vert Leap (in)		176,00	169,00	164,00	109,00	158,00	
Comb_Broad Jump (in)		169,00	173,00	161,00	108,00	154,00	
Comb_Shuttle		164,00	161,00	169,00	112,00	151,00	
Comb_3Cone		109,00	108,00	112,00	114,00	106,00	
College_G		158,00	154,00	151,00	106,00	202,00	
College_TD		158,00	154,00	151,00	106,00	202,00	
College_INT		158,00	154,00	151,00	106,00	202,00	
College_RAT		158,00	154,00	151,00	106,00	202,00	

Correlações

		College_TD	College_INT	College_RAT
Correlação de Pearson	Draft_Selected	-,16	,18	-,22
	Comb_Height (in)	-,09	-,09	-,04
	Comb_Weight (lbs)	,02	-,19	,02
	Comb_Wonderlic	,07	,02	,04
	Comb_40 Yard	,00	,20	-,11
	Comb_Vert Leap (in)	,06	-,11	,03
	Comb_Broad Jump (in)	-,01	-,17	,09
	Comb_Shuttle	-,05	,12	-,19

	Comb_3Cone	-,01	,13	-,22
	College_G	,56	,51	,34
	College_TD	1,00	,52	,42
	College_INT	,52	1,00	,04
	College_RAT	,42	,04	1,00
Sig. (1 extremidade)	Draft_Selected	,02	,01	,00
	Comb_Height (in)	,10	,09	,28
	Comb_Weight (lbs)	,36	,00	,39
	Comb_Wonderlic	,22	,42	,33
	Comb_40 Yard	,48	,00	,07
	Comb_Vert Leap (in)	,22	,09	,33
	Comb_Broad Jump (in)	,44	,02	,14
	Comb_Shuttle	,26	,08	,01
	Comb_3Cone	,47	,09	,01
	College_G	,00	,00	,00
	College_TD	.	,00	,00
	College_INT	,00	.	,29
	College_RAT	,00	,29	.
N	Draft_Selected	182,00	182,00	182,00
	Comb_Height (in)	202,00	202,00	202,00
	Comb_Weight (lbs)	202,00	202,00	202,00
	Comb_Wonderlic	114,00	114,00	114,00
	Comb_40 Yard	182,00	182,00	182,00
	Comb_Vert Leap (in)	158,00	158,00	158,00
	Comb_Broad Jump (in)	154,00	154,00	154,00
	Comb_Shuttle	151,00	151,00	151,00
	Comb_3Cone	106,00	106,00	106,00
	College_G	202,00	202,00	202,00
	College_TD	202,00	202,00	202,00
	College_INT	202,00	202,00	202,00
	College_RAT	202,00	202,00	202,00

Variáveis Inseridas/Removidas^a

Modelo	Variáveis inseridas	Variáveis removidas	Método
1	College_RAT, Comb_Weight (lbs), Comb_Wonderlic, College_INT, Comb_Vert Leap (in), Comb_3Cone, College_G, Comb_Height (in), Comb_Shuttle, College_TD, Comb_40 Yard, Comb_Broad Jump (in) ^b		Inserir

a. Variável Dependente: Draft_Selected

b. Todas as variáveis solicitadas inseridas.

Resumo do modelo^b

Modelo	R	R quadrado	R quadrado ajustado	Erro padrão da estimativa	Estatísticas de mudança	
					Mudança de R quadrado	Mudança F
1	,58 ^a	,34	,23	64,85	,34	3,33

Resumo do modelo^b

Modelo	Estatísticas de mudança			
	df1	df2	Sig. Mudança F	
1	12,00	79,00	,00	2,04

a. Preditores: (Constante), College_RAT, Comb_Weight (lbs), Comb_Wonderlic, College_INT, Comb_Vert Leap (in), Comb_3Cone, College_G, Comb_Height (in), Comb_Shuttle, College_TD, Comb_40 Yard, Comb_Broad Jump (in)

b. Variável Dependente: Draft_Selected

ANOVA^a

Modelo		Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
1	Regressão	167903,91	12,00	13991,99	3,33	,00 ^b
	Resíduo	332236,42	79,00	4205,52		
	Total	500140,33	91,00			

a. Variável Dependente: Draft_Selected

b. Preditores: (Constante), College_RAT, Comb_Weight (lbs), Comb_Wonderlic, College_INT, Comb_Vert Leap (in), Comb_3Cone, College_G, Comb_Height (in), Comb_Shuttle, College_TD, Comb_40 Yard, Comb_Broad Jump (in)

Coefficientes^a

Modelo		Coefficients não padronizados		Coefficientes padronizados	t	Sig.
		B	Erro Erro	Beta		
1	(Constante)	763,89	534,88		1,43	,16
	Comb_Height (in)	-9,95	5,79	-,22	-1,72	,09
	Comb_Weight (lbs)	-,06	,66	-,01	-,09	,92
	Comb_Wonderlic	-,04	1,15	,00	-,04	,97
	Comb_40 Yard	-9,33	60,33	-,02	-,15	,88
	Comb_Vert Leap (in)	-5,05	3,77	-,22	-1,34	,18
	Comb_Broad Jump (in)	-,35	1,66	-,03	-,21	,83
	Comb_Shuttle	-75,98	53,76	-,18	-1,41	,16
	Comb_3Cone	89,83	36,48	,33	2,46	,02
	College_G	1,86	,97	,23	1,91	,06
	College_TD	-,93	,36	-,35	-2,62	,01
	College_INT	1,06	,74	,18	1,44	,15
	College_RAT	-,32	,28	-,12	-1,13	,26

Coeficientes^a

Modelo		Correlações			Estatísticas de colinearidade	
		Ordem zero	Parcial	Parte	Tolerância	VIF
1	(Constante)					
	Comb_Height (in)	-,21	-,19	-,16	,52	1,93
	Comb_Weight (lbs)	-,27	-,01	-,01	,53	1,87
	Comb_Wonderlic	-,06	,00	,00	,76	1,31
	Comb_40 Yard	,22	-,02	-,01	,37	2,73
	Comb_Vert Leap (in)	-,28	-,15	-,12	,30	3,33
	Comb_Broad Jump (in)	-,30	-,02	-,02	,35	2,84
	Comb_Shuttle	,16	-,16	-,13	,52	1,91
	Comb_3Cone	,33	,27	,23	,48	2,09
	College_G	,10	,21	,18	,57	1,75
	College_TD	-,16	-,28	-,24	,48	2,07
	College_INT	,18	,16	,13	,54	1,86
	College_RAT	-,22	-,13	-,10	,69	1,44

a. Variável Dependente: Draft_Selected

Estatísticas de resíduos^a

	Mínimo	Máximo	Média	Erro Desvio	N
Valor previsto	-25,57	154,41	76,84	39,17	76,00
Erro Valor previsto	-2,62	1,57	-,24	,91	76,00
Erro padrão do valor previsto	15,77	86,50	24,11	8,75	76,00
Valor previsto ajustado	-29,97	153,14	72,45	40,75	69,00
Resíduo	-111,42	143,90	7,49	60,82	70,00
Erro Resíduo	-1,72	2,22	,12	,94	70,00
Resíduos Resíduo	-1,77	2,49	,13	1,02	69,00
de Estud.	-118,67	180,78	9,65	71,35	69,00
Resíduos de Estud.	-1,80	2,57	,13	1,03	69,00
Mahal. Distância	4,39	160,90	13,23	17,96	76,00
Distância de Cook	,00	,12	,01	,02	69,00
Valor de ponto alavanca centralizado	,05	1,77	,15	,20	76,00

a. Variável Dependente: Draft_Selected

GRÁFICOS

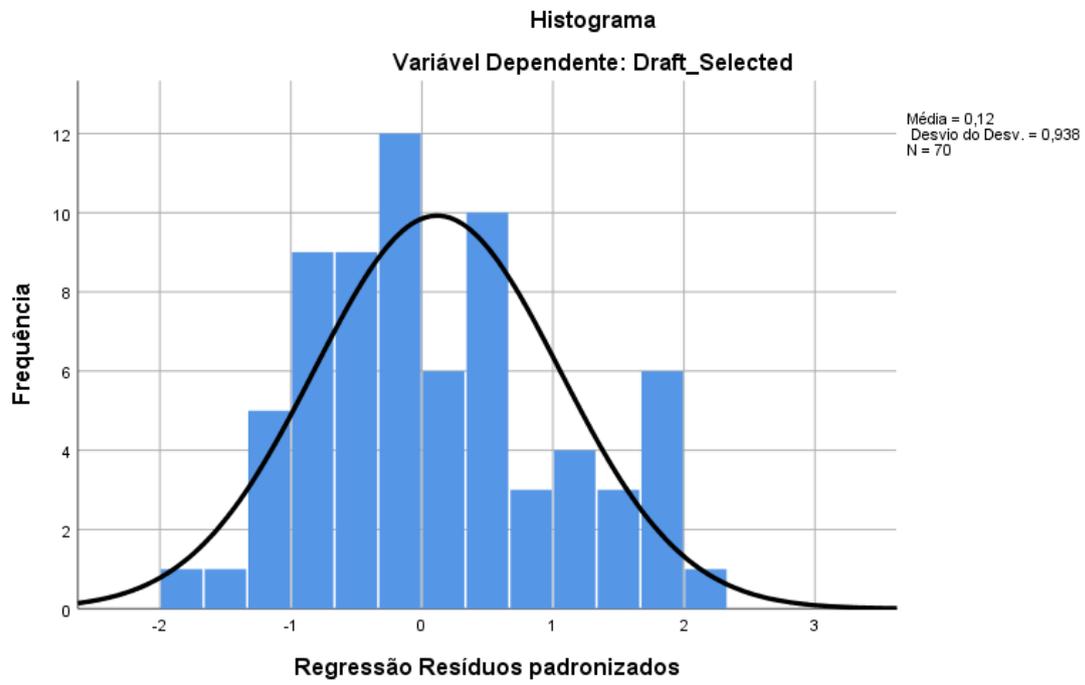


Gráfico P-P Normal de Regressão Resíduos padronizados

