

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS**  
**NÚCLEO DE EDUCAÇÃO A DISTÂNCIA**  
**Pós-graduação *Lato Sensu* em Ciência de Dados e Big Data**

**Gustavo Henrique de Lima Perencio**

**Análise de dados de lutas no evento esportivo**  
**Ultimate Fighting Championship**

Belo Horizonte  
2021

**Gustavo Henrique de Lima Perencio**

**ANÁLISE DE DADOS DE LUTAS NO EVENTO ESPORTIVO  
ULTIMATE FIGHTING CHAMPIONSHIP**

Trabalho de Conclusão de Curso apresentado  
ao Curso de Especialização em Ciência de  
Dados e Big Data como requisito parcial à  
obtenção do título de especialista.

Belo Horizonte

2021

## SUMÁRIO

<b>1. Introdução</b>	<b>4</b>
<b>1.1. Contextualização</b>	<b>4</b>
<b>1.1. O problema proposto</b>	<b>4</b>
<b>2. Coleta de Dados</b>	<b>4</b>
<b>3. Processamento/Tratamento de Dados</b>	<b>5</b>
<b>4. Análise e Exploração dos Dados</b>	<b>5</b>
<b>5. Criação de Modelos de Machine Learning</b>	<b>5</b>
<b>6. Apresentação dos Resultados</b>	<b>5</b>
<b>7. Links</b>	<b>6</b>
<b>REFERÊNCIAS</b>	<b>7</b>

## 1. Introdução

### 1.1. Contextualização

O presente trabalho propõe a análise e interpretação de dados de lutadores e lutas do evento esportivo de combate UFC (*Ultimate Fighting Championship*). Os dados são uma compilação encontrada no site [kaggle](#) sendo uma compilação de dados obtidos via scrapping no site oficial da organização esportiva, o [ufc fight stats](#). Propõe-se, também, um modelo de machine learning que visa de alguma forma a partir dos dados de lutas coletados prever em certo grau resultados futuros.

### 1.2. O problema proposto

O projeto descrito por este trabalho visa encontrar alguns padrões em dados de lutas feitas no evento esportivo UFC descritas entre os anos de 1993 e 2019 para calibrar um modelo de aprendizado de máquina que tem como objetivo apontar o vencedor de lutas dados certos tipos de dados de um atleta. Com isso, caso o modelo tenha uma eficácia satisfatória, pode ser usado como referência de palpites em sites de bolsa de apostas ou apenas para fins acadêmicos.

Os dados coletados foram disponibilizados na plataforma [kaggle](#) e foram obtidos por meio de *web crawlers* do site [ufc fight stats](#), uma ferramenta *online* que publica dados de lutas e atletas do UFC desde seu início no ano de 1993 e é atualizado após cada evento.

A análise visa buscar por padrões nos dados que possam de alguma forma se correlacionar e propor a partir destes dados um modelo de *machine learning* que consiga descrever com certo nível de acurácia resultados de lutas futuras as quais sejam inseridos dados compatíveis dos atletas.

O período dos dados analisados vem desde o começo do UFC em 1993 e possui informações de lutas até o ano de 2019. Ou seja, o modelo se baseará em 26 anos de lutas no campeonato.

## 2. Coleta de Dados

A análise possui três fontes de dados obtidas no site [kaggle](#) uma chamada *raw\_fighter\_details.csv*, que é uma descrição dos atributos dos(as) atletas do evento, *raw\_total\_fight\_data.csv*, que é uma descrição das lutas do evento e a outra *data.csv*, que é uma compilação das duas tabelas anteriores para fins de análise. Para o modelo foram usadas as tabelas *raw\_fighter\_details.csv* e *data.csv*. Na tabela abaixo estão descritas as colunas utilizadas de cada *dataset*

### *data.csv*:

Esta tabela é um consolidado dos dados do lutador do lado vermelho (melhor colocado no ranking) e lado azul (desafiante), além de resultado da luta, tipo de golpes aplicados, árbitro da partida entre outros. Os mesmos campos se aplicam para os lutadores do corner azul (B\_<nome\_do\_atributo>) e vermelho (R\_<nome\_do\_atributo>):

Nome da coluna/campo	Descrição	Tipo
B_fighter	nome do atleta do corner azul	string
Winner	Vencedor	string
weight_class	Categoria de peso da luta	string
B_current_lose_streak	Número de derrotas em sequência do lutador do corner azul	float
B_current_win_streak	Número de vitórias em sequência do lutador do corner azul	float
B_draw	Número de empates do lutador do corner azul	float
B_avg_BODY_att	Média de ataques no corpo do lutador do corner azul	float
B_avg_BODY_landed	Média de ataques eficazes no corpo desferidos pelo lutador do corner azul	float
B_avg_CLINCH_att	Média de ataques desferidos do clinch do lutador do corner azul	float
B_avg_CLINCH_landed	Média de ataques eficazes desferidos do clinch do lutador do corner azul	float

B_avg_DISTANCE_att	Média de ataques a distância desferidos pelo lutador do corner azul	float
B_avg_DISTANCE_landed	Média de ataques a distância eficazes desferidos pelo lutador do corner azul	float
B_avg_GROUND_att	Média de ataques no solo desferidos pelo lutador do corner azul	float
B_avg_GROUND_landed	Média de ataques eficazes no solo desferidos pelo lutador do corner azul	float
B_avg_HEAD_att	Média de ataques na cabeça desferidos pelo lutador do corner azul	float
B_avg_HEAD_landed	Média de ataques eficazes na cabeça desferidos pelo lutador do corner azul	float
B_avg_KD	Média de knockdowns desferidos pelo lutador do corner azul	float
B_avg_LEG_att	Média de ataques nas pernas desferidos pelo lutador do corner azul	float
B_avg_LEG_landed	Média de ataques eficazes nas pernas desferidos pelo lutador do corner azul	float
B_avg_PASS	Média de passagens de guarda desferidos pelo lutador do corner azul	float
B_avg_REV	Média de reversões desferidos pelo lutador do corner azul	float
B_avg_SIG_STR_att	Média de ataques significativos desferidos pelo lutador do corner azul	float
B_avg_SIG_STR_landed	Média de ataques significativos eficazes desferidos pelo lutador do corner azul	float
B_avg_SIG_STR_pct	Média de ataques significativos eficazes desferidos pelo lutador do corner azul em %	float
B_avg_SUB_ATT	Média de tentativas de submissão desferidas pelo lutador do corner azul	float
B_avg_TD_att	Média de tentativas de queda desferidas pelo lutador do corner azul	float

B_avg_TD_landed	Média de quedas eficazes desferidas pelo lutador do corner azul	float
B_avg_TD_pct	Média de quedas desferidas pelo lutador do corner azul em %	float
B_avg_TOTAL_STR_att	Média total de ataques desferidos pelo lutador do corner azul	float
B_avg_TOTAL_STR_landed	Média total de ataques eficazes desferidos pelo lutador do corner azul	float
B_longest_win_streak	sequência mais longa de vitórias do lutador do corner azul	float
B_losses	Número de derrotas do lutador do corner azul	float
B_avg_opp_BODY_att	Média de ataques do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_BODY_landed	Média de ataques eficazes do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_CLINCH_att	Média de ataques no clinch do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_CLINCH_landed	Média de ataques eficazes no clinch do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_DISTANCE_att	Média de ataques à distância do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_DISTANCE_landed	Média de ataques eficazes à distância do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_GROUND_att	Média de ataques no solo do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_GROUND_landed	Média de ataques eficazes no solo do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_HEAD_att	Média de ataques na cabeça do oponente desferidos contra o lutador do corner azul	float

B_avg_opp_HEAD_landed	Média de ataques eficazes na cabeça do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_KD	Média de quedas do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_LEG_att	Média de ataques nas pernas do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_LEG_landed	Média de ataques eficazes nas pernas do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_PASS	Média de passagens de guarda do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_REV	Média de reversões do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_SIG_STR_att	Média de ataques significativos do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_SIG_STR_landed	Média de ataques significativos eficazes do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_SIG_STR_pct	Média de ataques significativos percentual do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_SUB_ATT	Média de submissões do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_TD_att	Média de tentativas de quedas do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_TD_landed	Média de tentativas de quedas eficazes do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_TD_pct	Média de tentativas de quedas percentual do oponente desferidos contra o lutador do corner azul	float



B_avg_opp_TOTAL_STR_att	Média total de ataques do oponente desferidos contra o lutador do corner azul	float
B_avg_opp_TOTAL_STR_Ianded	Média total de ataques eficazes do oponente desferidos contra o lutador do corner azul	float
B_total_rounds_fought	Número de rounds lutados pelo lutador do corner azul	float
B_total_time_fought(seconds)	Tempo total lutado pelo lutador do corner azul	float
B_total_title_bouts	Número de disputas de título pelo lutador do corner azul	float
B_win_by_Decision_Majority	Número de vitórias por decisão majoritária pelo lutador do corner azul	float
B_win_by_Decision_Split	Número de vitórias por decisão dividida pelo lutador do corner azul	float
B_win_by_Decision_Unanimous	Número de vitórias por decisão unanime pelo lutador do corner azul	float
B_win_by_KO/TKO	Número de vitórias por knockout ou knockout técnico do lutador do corner azul	float
B_win_by_Submission	Número de vitórias por submissão do lutador do corner azul	float
B_win_by_TKO_Doctor_Stoppage	Número de vitórias por knockout técnico devido a parada médica do lutador do corner azul	float
B_wins	Número de vitórias do lutador do corner azul	float
B_Height_cms	Altura do lutador do corner azul	float
B_Reach_cms	Alcance do lutador do corner azul	float
B_Weight_lbs	Peso do lutador do corner azul	float
B_age	Idade do lutador do corner azul	float
...		
Mesmos campos se repetem para os lutadores do corner vermelho		

*raw\_fighter\_details.csv:*

Nome da coluna/campo	Descrição	Tipo
fighter_name	nome do atleta	string
Height	Altura do atleta	float
Weight	Peso do atleta	float
Reach	Alcance do atleta	float
DOB	Data de nascimento	string
TODAY	Coluna customizada com a data de hoje	Date
Age	Coluna customizada com a idade do atleta Resultado de: (col['TODAY'] - col['DOB'])	int

### 3. Processamento/Tratamento de Dados

A tabela `data.cvc` foi processada desconsiderando-se as colunas ["Referee", "date", "location", "title\_bout", "no\_of\_rounds", "R\_Stance", "B\_Stance"] pois essas não possuem dados que iriam contribuir com a análise portanto sua extração simplificou o processamento dos dados. Este *dataframe* possui 5144 registros de lutas no evento e sendo que nenhum destes é duplicado. O número de valores null ou *not a number* (NaN) é descrito pela tabela abaixo. Estes valores foram substituídos por zero no tratamento dos dados para que as equações de machine learning não fossem feitas em NaN deixando sua acurácia mais baixa.

R_fighter	0
B_fighter	0
Winner	0
weight_class	0
B_current_lose_streak	0
B_current_win_streak	0
B_draw	0
B_avg_BODY_att	1265
B_avg_BODY_landed	1265
B_avg_CLINCH_att	1265
B_avg_CLINCH_landed	1265

B_avg_DISTANCE_att	1265
B_avg_DISTANCE_landed	1265
B_avg_GROUND_att	1265
B_avg_GROUND_landed	1265
B_avg_HEAD_att	1265
B_avg_HEAD_landed	1265
B_avg_KD	1265
B_avg_LEG_att	1265
B_avg_LEG_landed	1265
B_avg_PASS	1265
B_avg_REV	1265
B_avg_SIG_STR_att	1265
B_avg_SIG_STR_landed	1265
B_avg_SIG_STR_pct	1265
B_avg_SUB_ATT	1265
B_avg_TD_att	1265
B_avg_TD_landed	1265
B_avg_TD_pct	1265
B_avg_TOTAL_STR_att	1265
B_avg_TOTAL_STR_landed	1265
B_longest_win_streak	0
B_losses	0
B_avg_opp_BODY_att	1265
B_avg_opp_BODY_landed	1265
B_avg_opp_CLINCH_att	1265
B_avg_opp_CLINCH_landed	1265
B_avg_opp_DISTANCE_att	1265
B_avg_opp_DISTANCE_landed	1265
B_avg_opp_GROUND_att	1265
B_avg_opp_GROUND_landed	1265
B_avg_opp_HEAD_att	1265
B_avg_opp_HEAD_landed	1265
B_avg_opp_KD	1265
B_avg_opp_LEG_att	1265
B_avg_opp_LEG_landed	1265
B_avg_opp_PASS	1265
B_avg_opp_REV	1265

B_avg_opp_SIG_STR_att	1265
B_avg_opp_SIG_STR_landed	1265
B_avg_opp_SIG_STR_pct	1265
B_avg_opp_SUB_ATT	1265
B_avg_opp_TD_att	1265
B_avg_opp_TD_landed	1265
B_avg_opp_TD_pct	1265
B_avg_opp_TOTAL_STR_att	1265
B_avg_opp_TOTAL_STR_landed	1265
B_total_rounds_fought	0
B_total_time_fought(seconds)	1265
B_total_title_bouts	0
B_win_by_Decision_Majority	0
B_win_by_Decision_Split	0
B_win_by_Decision_Unanimous	0
B_win_by_KO/TKO	0
B_win_by_Submission	0
B_win_by_TKO_Doctor_Stoppage	0
B_wins	0
B_Height_cms	8
B_Reach_cms	666
B_Weight_lbs	6
R_current_lose_streak	0
R_current_win_streak	0
R_draw	0
R_avg_BODY_att	650
R_avg_BODY_landed	650
R_avg_CLINCH_att	650
R_avg_CLINCH_landed	650
R_avg_DISTANCE_att	650
R_avg_DISTANCE_landed	650
R_avg_GROUND_att	650
R_avg_GROUND_landed	650
R_avg_HEAD_att	650
R_avg_HEAD_landed	650
R_avg_KD	650
R_avg_LEG_att	650

R_avg_LEG_landed	650
R_avg_PASS	650
R_avg_REV	650
R_avg_SIG_STR_att	650
R_avg_SIG_STR_landed	650
R_avg_SIG_STR_pct	650
R_avg_SUB_ATT	650
R_avg_TD_att	650
R_avg_TD_landed	650
R_avg_TD_pct	650
R_avg_TOTAL_STR_att	650
R_avg_TOTAL_STR_landed	650
R_longest_win_streak	0
R_losses	0
R_avg_opp_BODY_att	650
R_avg_opp_BODY_landed	650
R_avg_opp_CLINCH_att	650
R_avg_opp_CLINCH_landed	650
R_avg_opp_DISTANCE_att	650
R_avg_opp_DISTANCE_landed	650
R_avg_opp_GROUND_att	650
R_avg_opp_GROUND_landed	650
R_avg_opp_HEAD_att	650
R_avg_opp_HEAD_landed	650
R_avg_opp_KD	650
R_avg_opp_LEG_att	650
R_avg_opp_LEG_landed	650
R_avg_opp_PASS	650
R_avg_opp_REV	650
R_avg_opp_SIG_STR_att	650
R_avg_opp_SIG_STR_landed	650
R_avg_opp_SIG_STR_pct	650
R_avg_opp_SUB_ATT	650
R_avg_opp_TD_att	650
R_avg_opp_TD_landed	650
R_avg_opp_TD_pct	650
R_avg_opp_TOTAL_STR_att	650

R_avg_opp_TOTAL_STR_landed	650
R_total_rounds_fought	0
R_total_time_fought(seconds)	650
R_total_title_bouts	0
R_win_by_Decision_Majority	0
R_win_by_Decision_Split	0
R_win_by_Decision_Unanimous	0
R_win_by_KO/TKO	0
R_win_by_Submission	0
R_win_by_TKO_Doctor_Stoppage	0
R_wins	0
R_Height_cms	4
R_Reach_cms	316
R_Weight_lbs	3
B_age	172
R_age	64

A coluna `weight_class` teve os valores NaN convertidos para 'NA' pois essa coluna é convertida no modelo como categórica sendo aplicado o algoritmo oneHot onde ela é convertida em 5 novas colunas com os pesos em lugar das strings.

A coluna `Winner` possui valores Blue, Red ou Draw, para o efeito do modelo essa coluna foi transformada em inteiros onde 0 significa que o combatente do corner azul ganhou, no caso de vitória do vermelho é colocado 1 e se deu empate -1.

O dataframe *raw\_fighter\_details.csv* possui 3313 entradas únicas dentre as quais possuem valores nulos ou NaN conforme a tabela a seguir:

fighter_name	0
Height	263
Weight	75
Reach	1744
DOB	740

Como este dataset foi usado apenas para demonstrar características de altura, peso, alcance e idade dos atletas as entradas com valores nulos foram desprezadas assim como a coluna *Stance* que possui dados sobre o tipo de postura principal do atleta e não seria utilizada neste contexto. Além disso, a coluna DOB foi convertida para datetime pois era string e possuía dados de nascimento dos atletas. Foi criada a coluna TODAY que guarda os dados do dia atual para ser utilizada como parâmetro de criação da coluna AGE que é o resultado da subtração do dia de hoje e a data de nascimento do lutador.

A coluna Height é ofertada no dataframe com as unidades do sistema inglês (Pés e polegadas) foi criada uma função de utilidade que recebe um valor em sistema inglês e devolve o resultado em centímetros. Todos os valores contidos nesta coluna foram atualizados segundo esse critério.

A coluna Reach é ofertada no dataframe com as unidades do sistema inglês (polegadas) e também teve seus valores atualizados para o sistema métrico.

A coluna Weight é dada em com as unidades do sistema inglês de peso (Libras) foi criada uma função de utilidade que dado um peso em libras devolve seu correspondente em Kg.

Com essas mudanças os dados estão preparados para análise.

#### **4. Análise e Exploração dos Dados**

Na análise dos dados ficaram evidentes que as categorias de peso mais ativas do UFC são peso Leve (Lightweight) com 19%, peso Meio-Médio (Welterweight) 18%, e peso Médio (Middleweight) com 14% das lutas. Isso pode estar relacionado a outro dado retirado do data frame que é o de *knockouts* nesse contexto as categorias mais pesadas (Heavyweight e Light Heavyweight) dispararam com médias próximas de dois o que são finalizações de lutas mais sérias demandando mais tempo de recuperação dos atletas o que pode identificar o porquê das categorias mais leves terem mais atividade.

Outra correlação a respeito disso é que nas 10 primeiras posições de lutadores com mais vitórias temos 9 lutadores que se encaixam nas categorias mais leves e apenas o lutador Frank Mir da categoria dos pesados.

Outro dado interessante retirado do dataset é que a média de idade da maioria absoluta de lutadores(as) está de 25 a 40 anos de idade. Dando uma janela de aproximadamente 15 anos de carreira para um atleta.

A altura média dos lutadores vai de 1 metro e 65 centímetros e 1 metro e 90 centímetros tendo alguns outliers fora deste padrão

O alcance médio dos lutadores está localizado entre 1 metro e 70 centímetros e 2 metros de cobertura.

## 5. Criação de Modelos de Machine Learning

Para o projeto foi desenvolvido um script em python para criar um modelo de Machine Learning baseado no algoritmo de classificação random forest. Este algoritmo foi selecionado por ser uma implementação bastante eficaz e difundida do modelo de classificação por árvores de decisão que utiliza seleção aleatória de features, escolhendo features adequadas a posição de nó raiz da árvore criando uma quantidade diversa de árvores o inibe a perda de generalização do modelo de árvore de decisão tradicional o chamado overfitting.

O data frame foi dividido em 75% de dados de treino e 25% de dados de teste, gerando uma acurácia final de aproximadamente 70%. O que diz que o modelo consegue prever 7 de 10 lutas fornecidas.

O trecho do código desenvolvido para o modelo é:

```
data = np.asarray(data)
X = data[:, :-1].astype(int)
y = data[:, -1].astype(int)
X_train, X_test, y_train, y_test =
train_test_split(X, y, test_size=.25, random_state=42)
clf = RandomForestClassifier(n_estimators=100)
model = clf.fit(X_train, y_train)
pred = model.predict(X_test)

# Accuracy: 0.6959564541213064
```

O código completo pode ser encontrado no google colab notebook hospedado no github [neste endereço](#).



## 6. Apresentação dos Resultados

### The Machine Learning Canvas

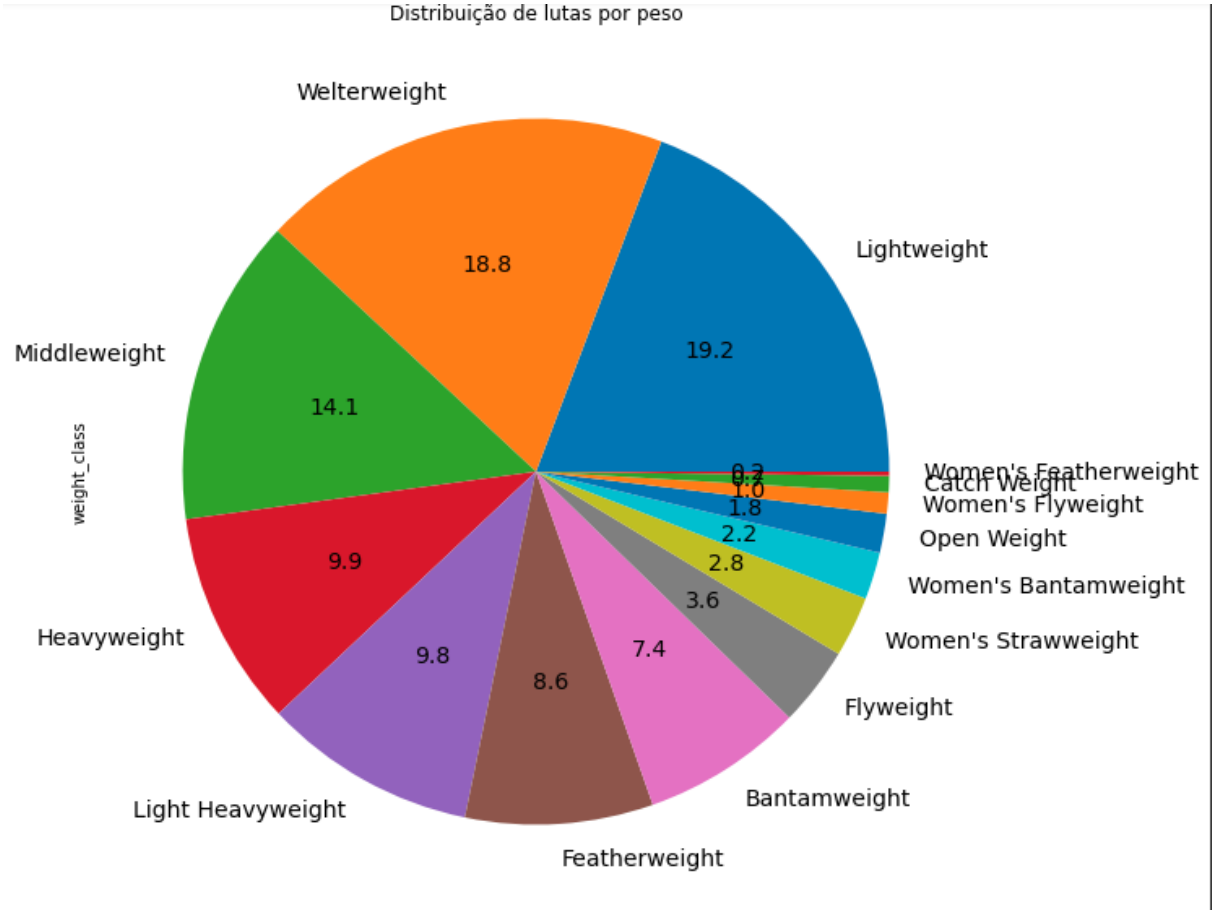
Date: Iteration: 2021

Designed for: TCC Big data Designed by: Gustavo Perencio

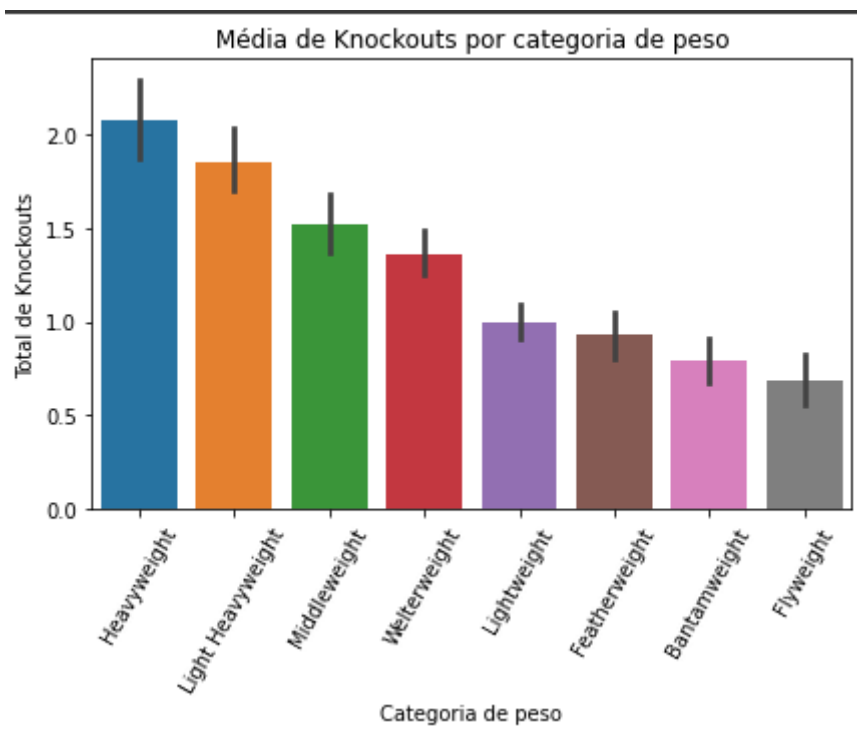
<b>Decisions</b>  How are predictions used to make decisions that provide the proposed value to the end-user?  Esse modelo visa analisar os dados de lutas do UFC para criar uma ferramenta de predição de resultados a partir dos dados de atletas	<b>ML task</b>  Input, output to predict, type of problem.  Inputs são os dados dos atletas que são fornecidos ao algoritmo de classificação random forest e este classifica o lutador como ganhador ou perdedor.	<b>Value Propositions</b>  What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving?  O valor a ser gerado é de servir de ferramenta para ajuda de decisão em qual atleta apostar em sistemas de aposta online baseado nos dados obtidos pelo modelo de aprendizado	<b>Data Sources</b>  Which raw data sources can we use (internal and external)?  Os dados foram extraídos do site <a href="#">kaggle</a> e esses por sua vez foram extraídos do site oficial da organização <a href="#">UFC stats</a>	<b>Collecting Data</b>  How do we get new data to learn from (inputs and outputs)?  No próprio link do <a href="#">UFC stats</a> na seção de lutadores pode ser feito um crawler para extrair dados e a partir deles criar novos casamentos de luta e expô-los ao modelo.
<b>Making Predictions</b>  When do we make predictions on new inputs? How long do we have to featurize a new input and make a prediction?  Sempre que haja uma nova luta marcada de atletas ainda não contemplados no modelo	<b>Offline Evaluation</b>  Methods and metrics to evaluate the system before deployment.  Foi feito o levantamento de acurácia do modelo por meio da função <code>accuracy_score</code> e este obteve um resultado de aproximadamente 70% o que inicialmente considerou-se como satisfatório.		<b>Features</b>  Input representations extracted from raw data sources.  As features são um conjunto de dados dos atletas como: Número de vitórias, golpes aplicados e suas efetividades, tipos de golpes, altura, peso, alcance, categoria de peso disputada, número de vitórias em sequência, número de derrotas e derrotas em sequência entre outros.	<b>Building Models</b>  When do we create/update models with new training data? How long do we have to featurize training inputs and create a model?  A princípio o modelo encontra-se criado.
	<b>Live Evaluation and Monitoring</b>  Methods and metrics to evaluate the system after deployment, and to quantify value creation.			

Dados gráficos extraídos na análise:

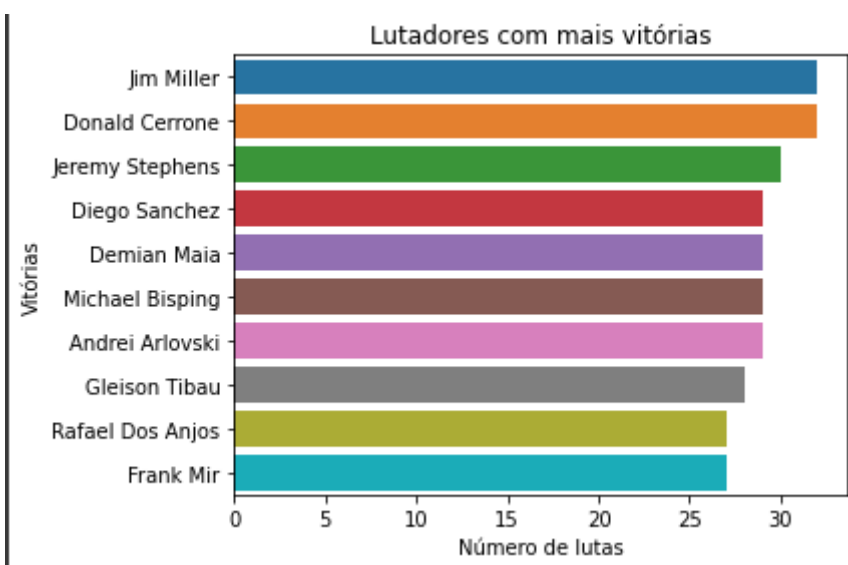
O gráfico abaixo apresenta o percentual de lutas por peso com as categorias Lightweight, Welterweight e Middleweight sendo as mais ativas da organização.



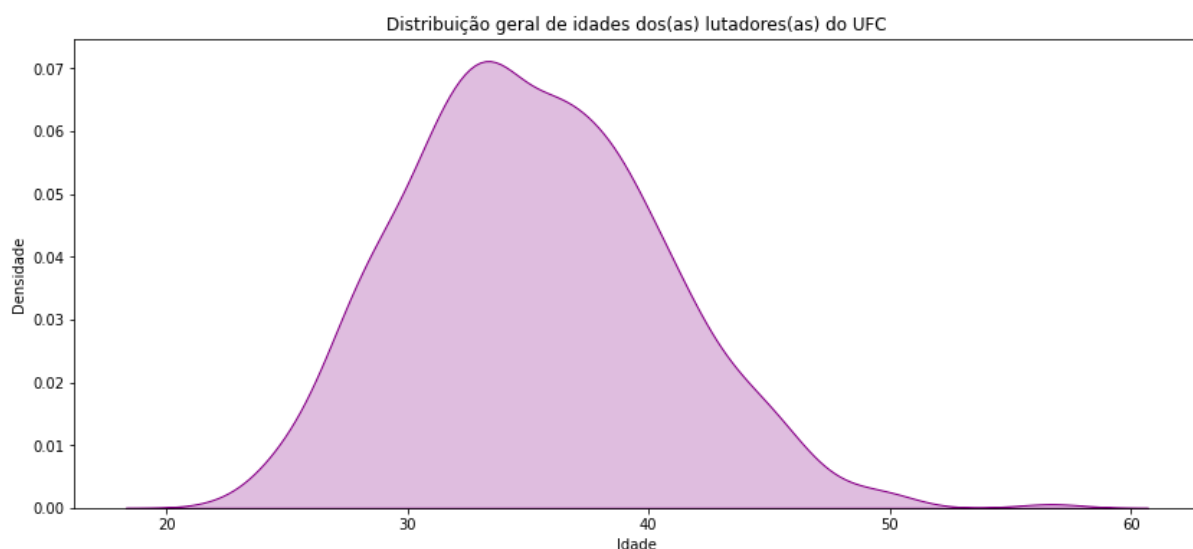
O gráfico abaixo demonstra a média de knockouts por categoria de peso tendo as categorias mais pesadas, Heavyweight e Light Heavyweight como as que mais tem esse tipo de finalização. Isso pode estar relacionado ao fato das categorias mais leves possuírem mais atividade pois knockouts são finalizações que demandam mais tempo de recuperação dos atletas.



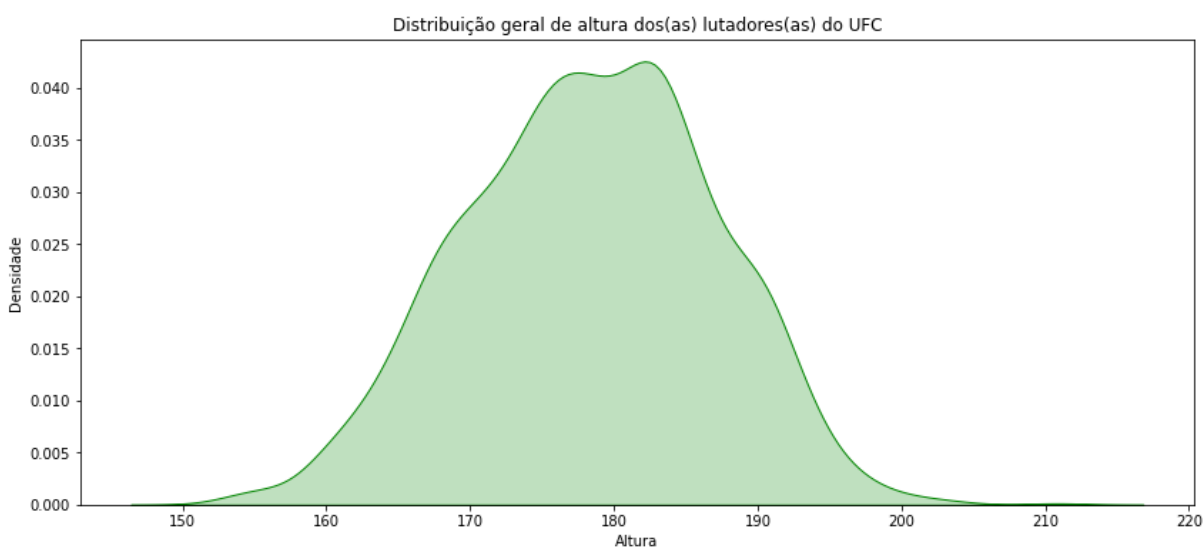
O gráfico abaixo é a relação dos atletas com mais vitórias no evento que estão relacionados no dataset. Esse gráfico mostra que em sua maioria os mais vitoriosos são os lutadores de categorias mais leves devido a essas lutas serem mais recorrentes.



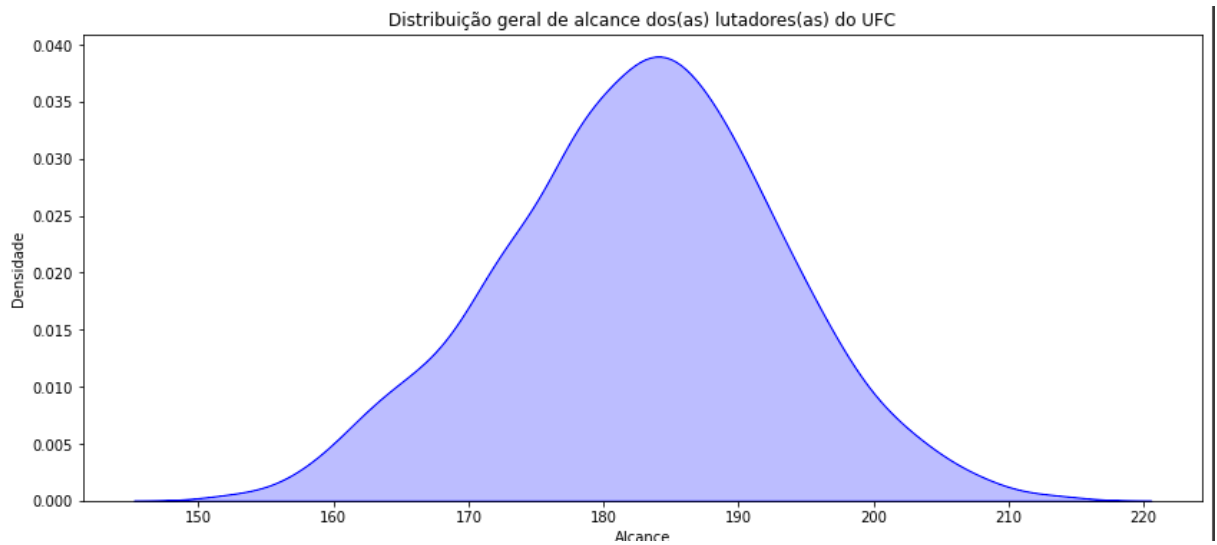
O gráfico abaixo mostra a distribuição média de idades dos lutadores do UFC. Nota-se que em sua grande maioria os lutadores estão entre 25 e 40 anos.



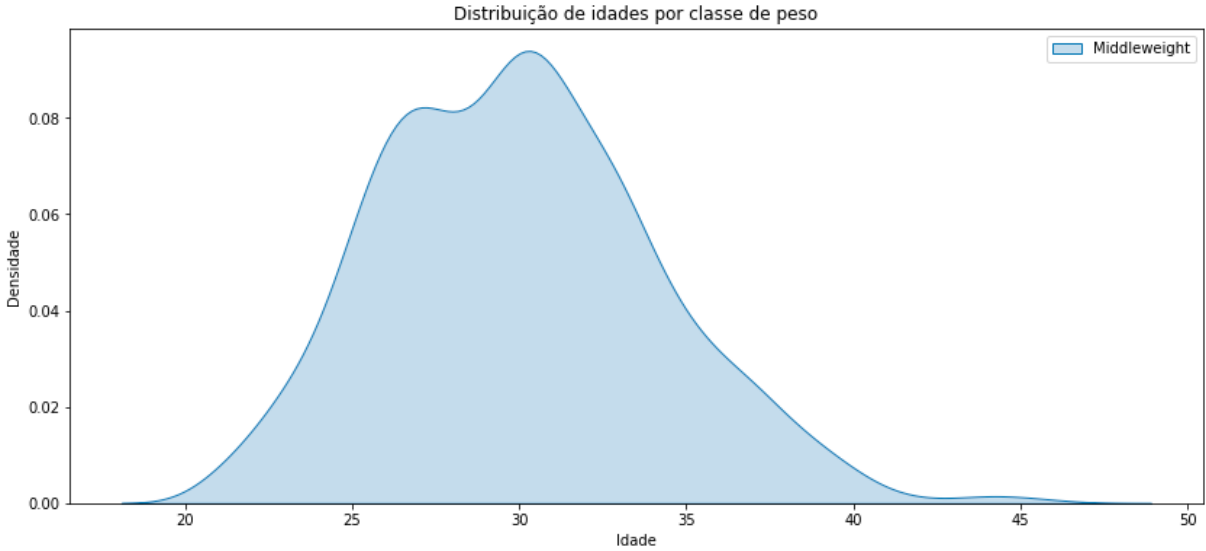
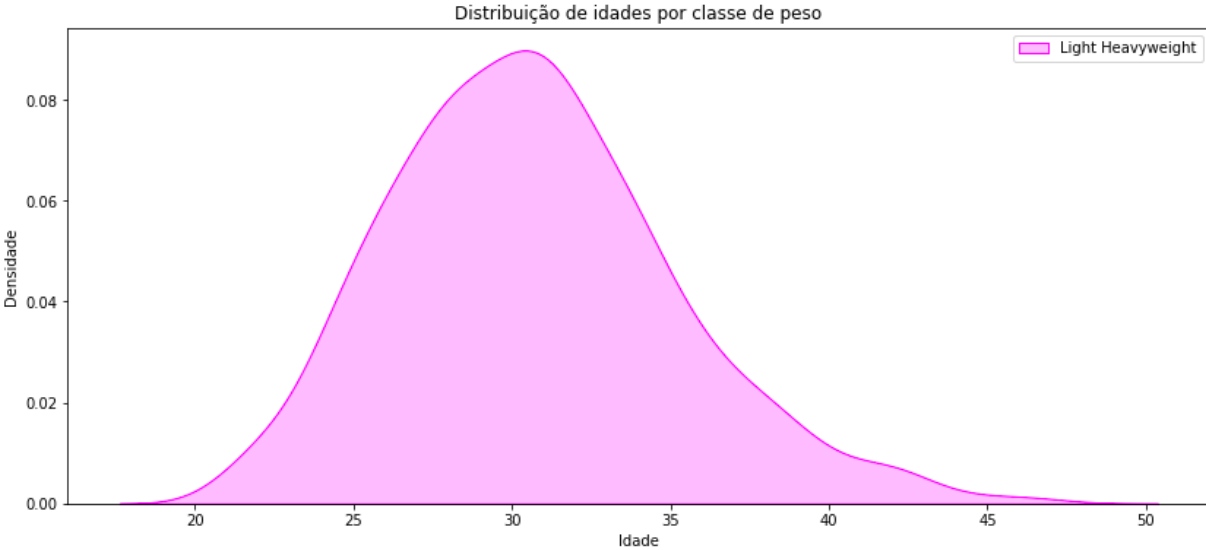
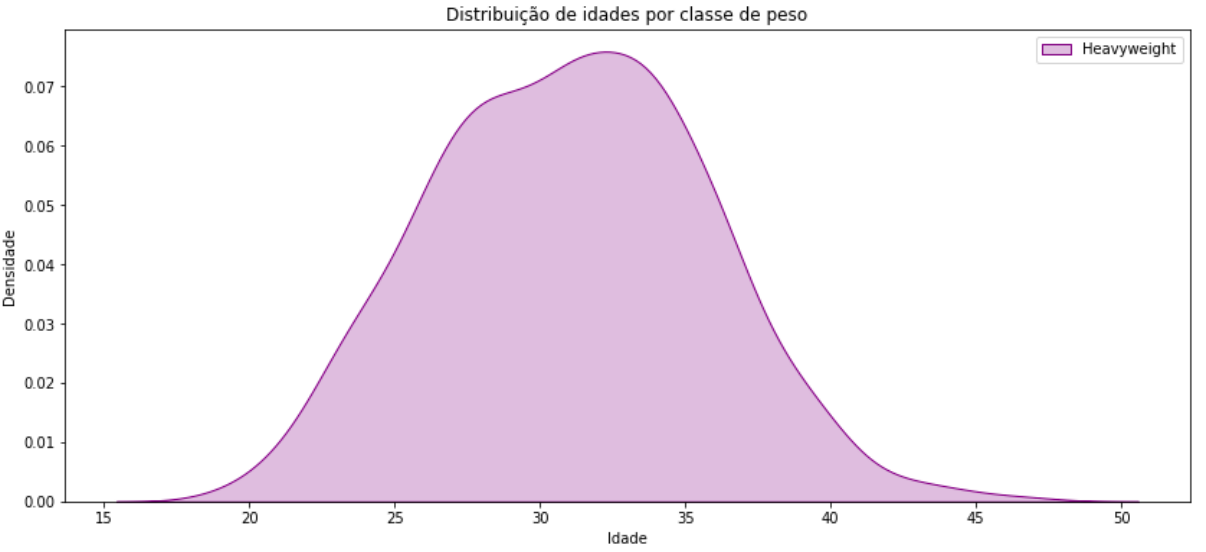
O gráfico a seguir mostra que os lutadores da organização tem uma altura em média entre 1,65m e 1,90m

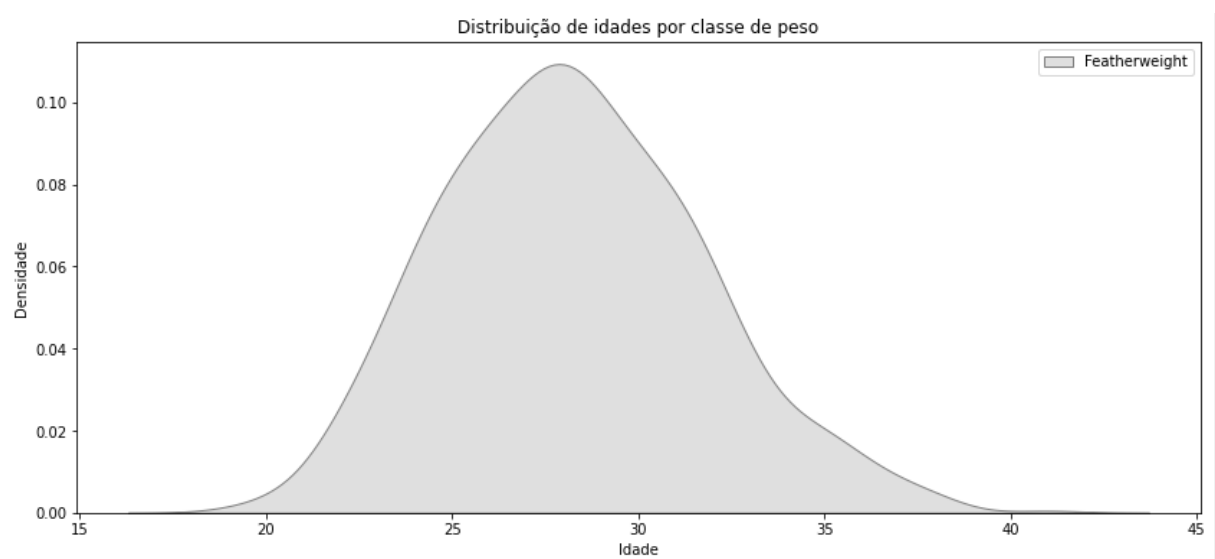
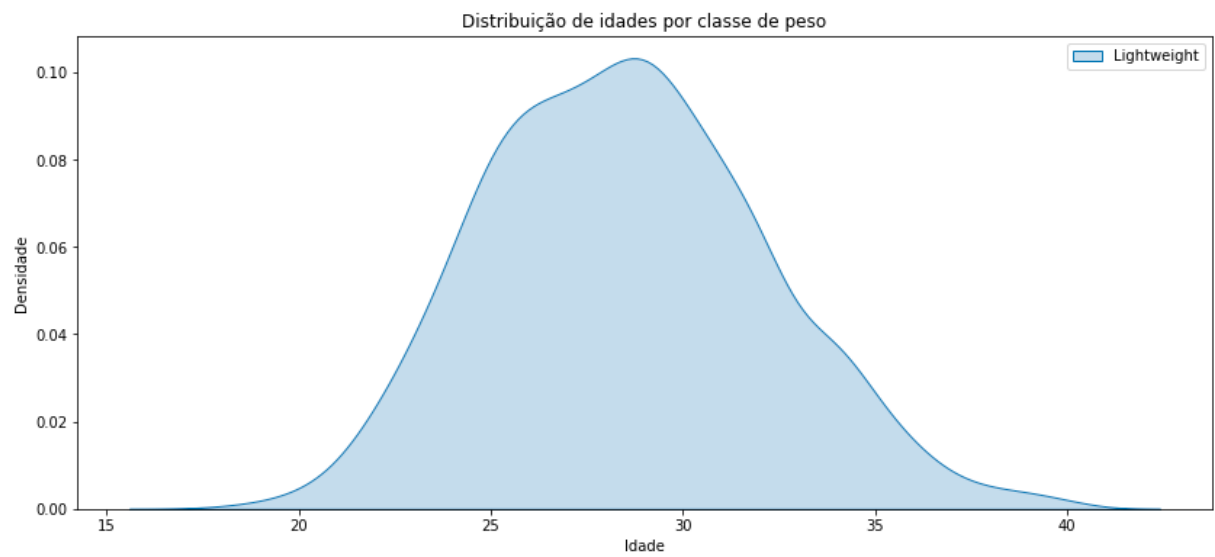
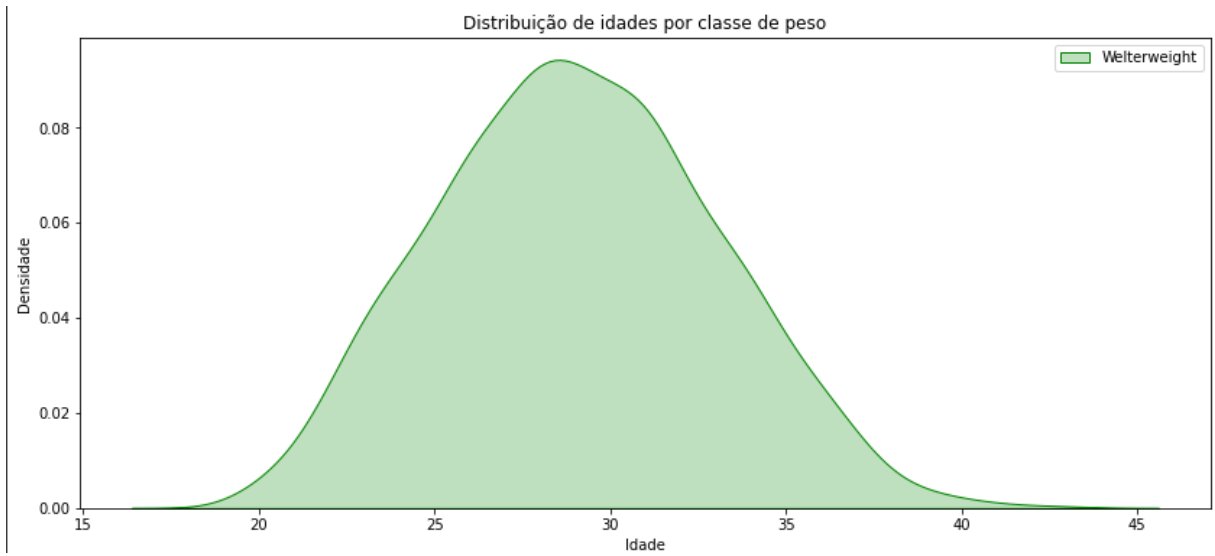


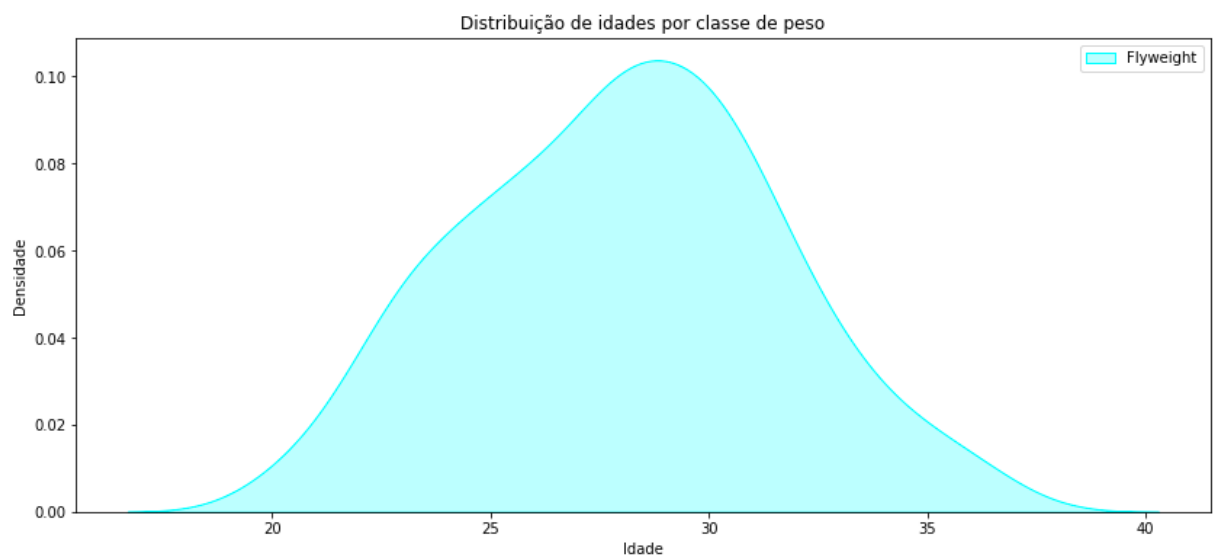
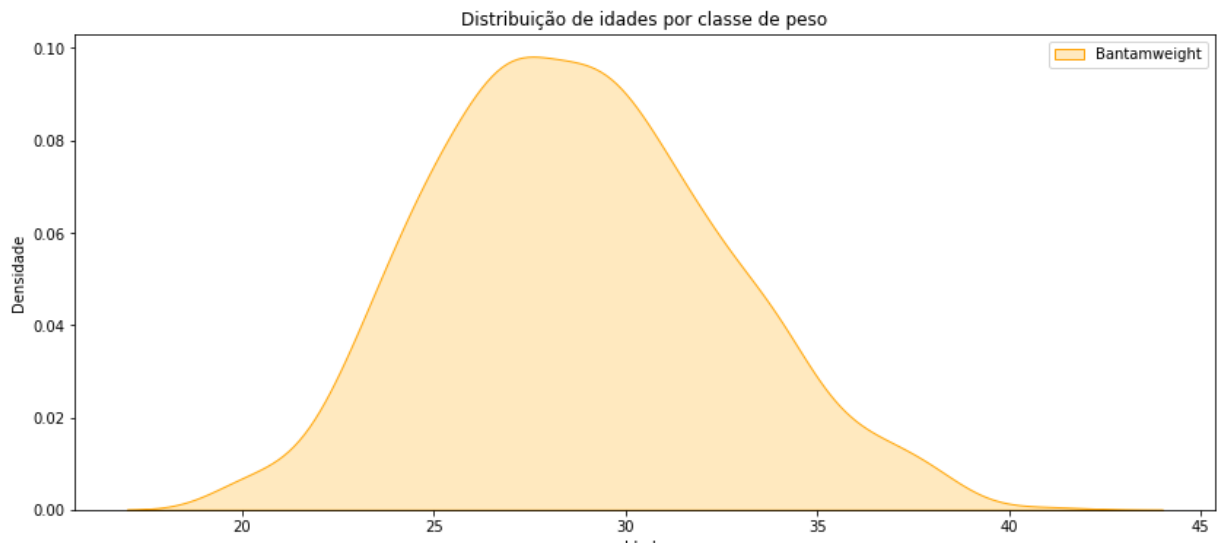
O próximo gráfico demonstra o alcance médio dos lutadores do UFC este concentra-se em média de 1,70m e 2m



Os gráficos abaixo demonstram as distribuições de idade organizadas por categorias de peso. Os gráficos demonstram que a categoria mais pesada é a que contribui mais para o aumento da média de idade sendo a que possui mais pessoas de idades maiores

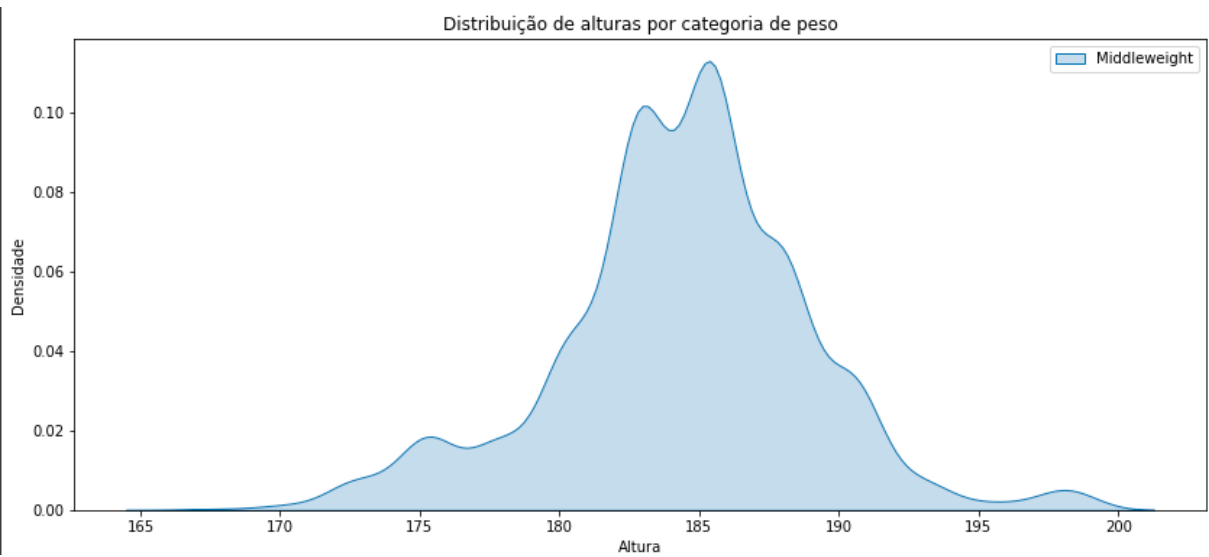
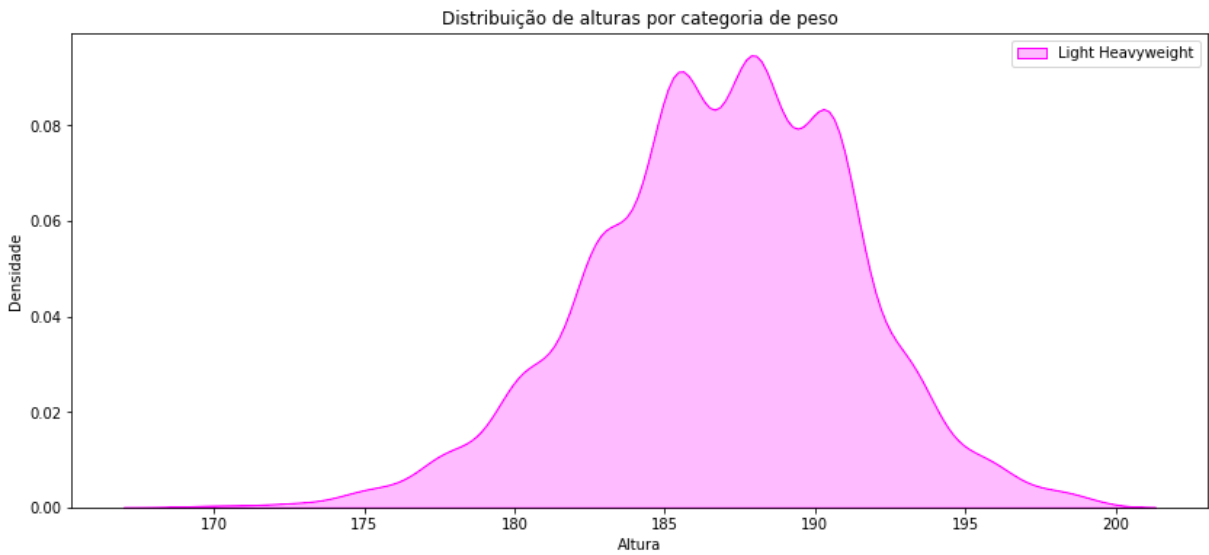
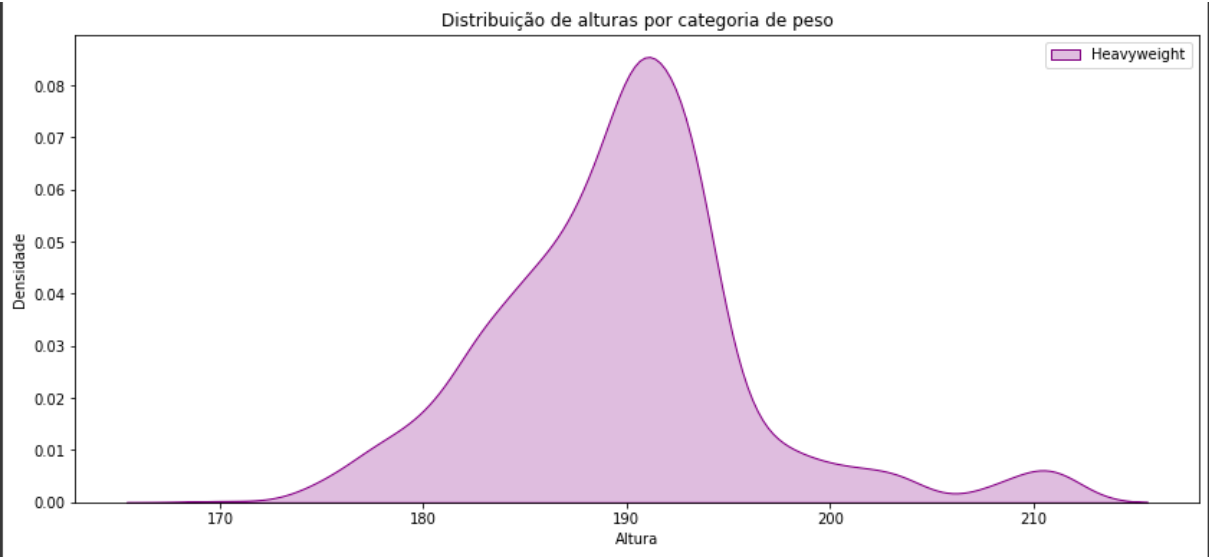


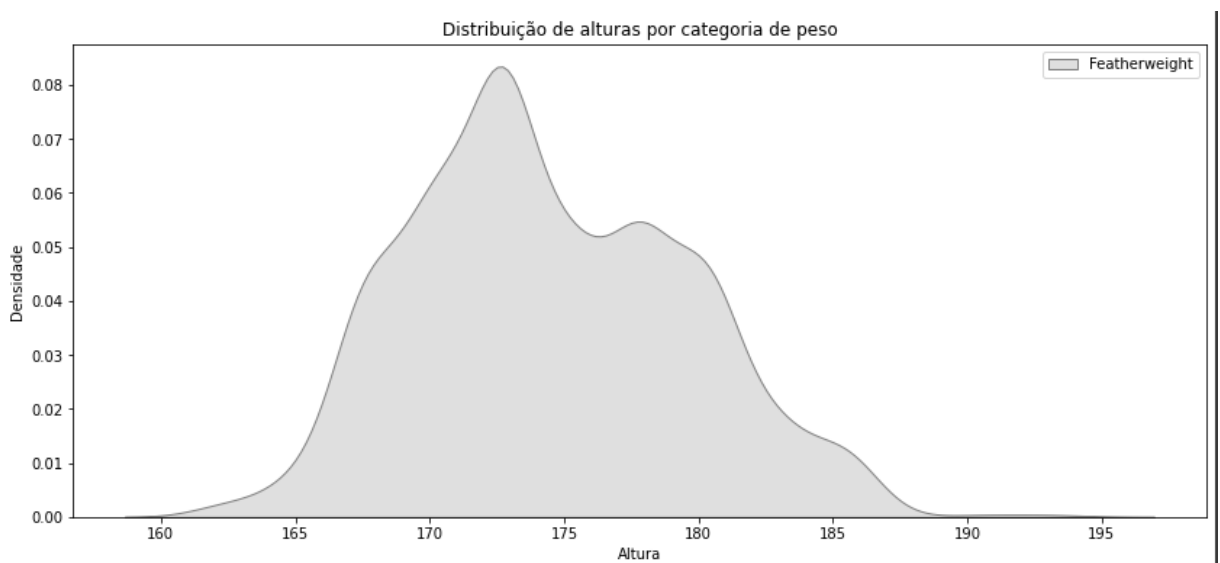
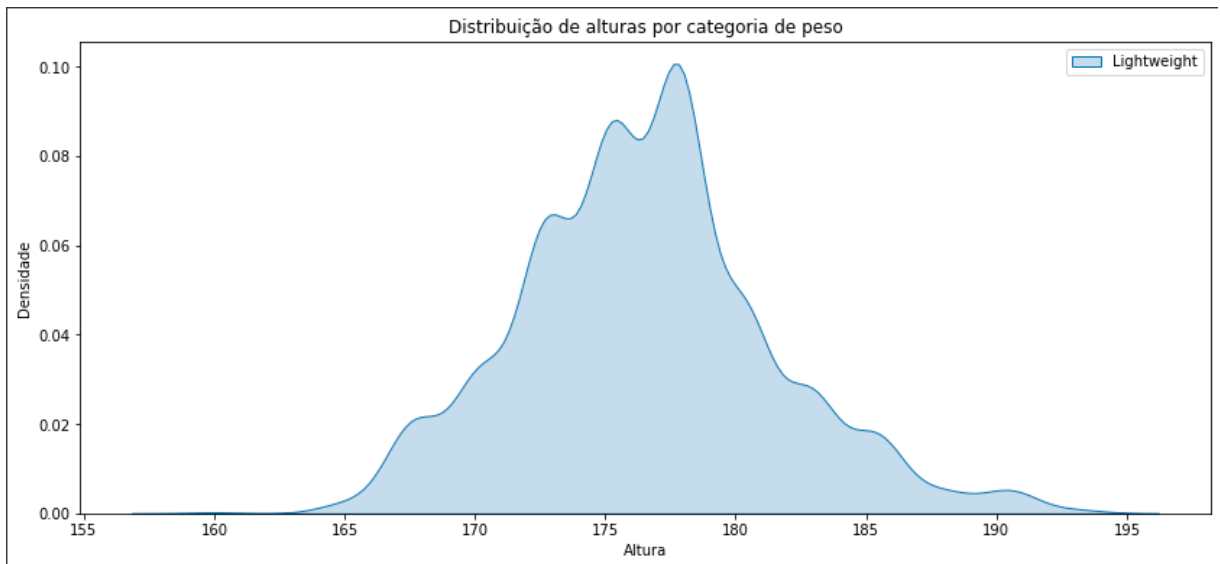
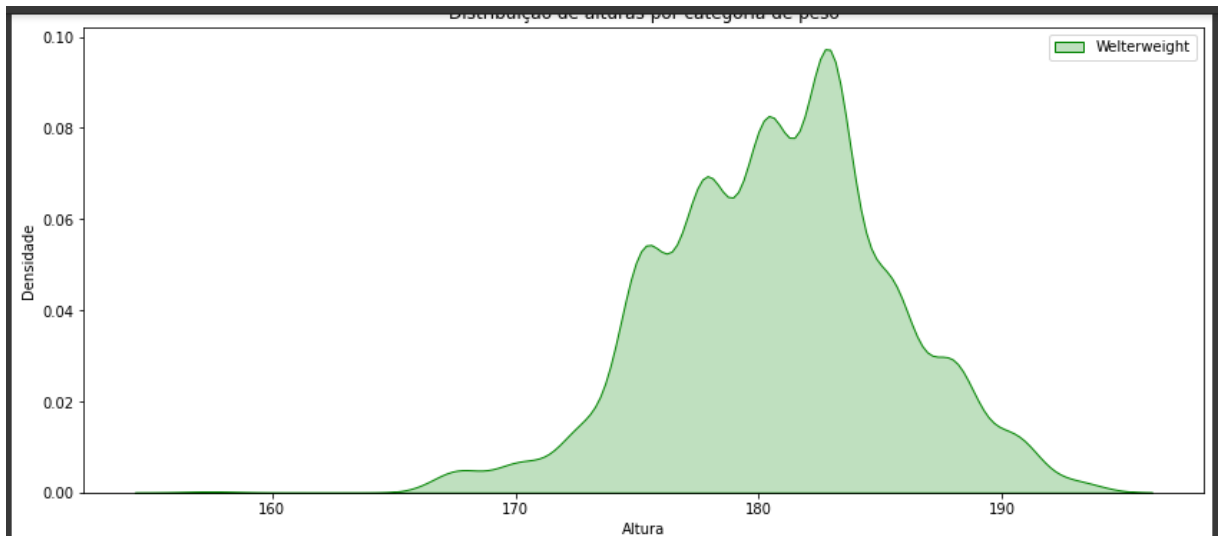


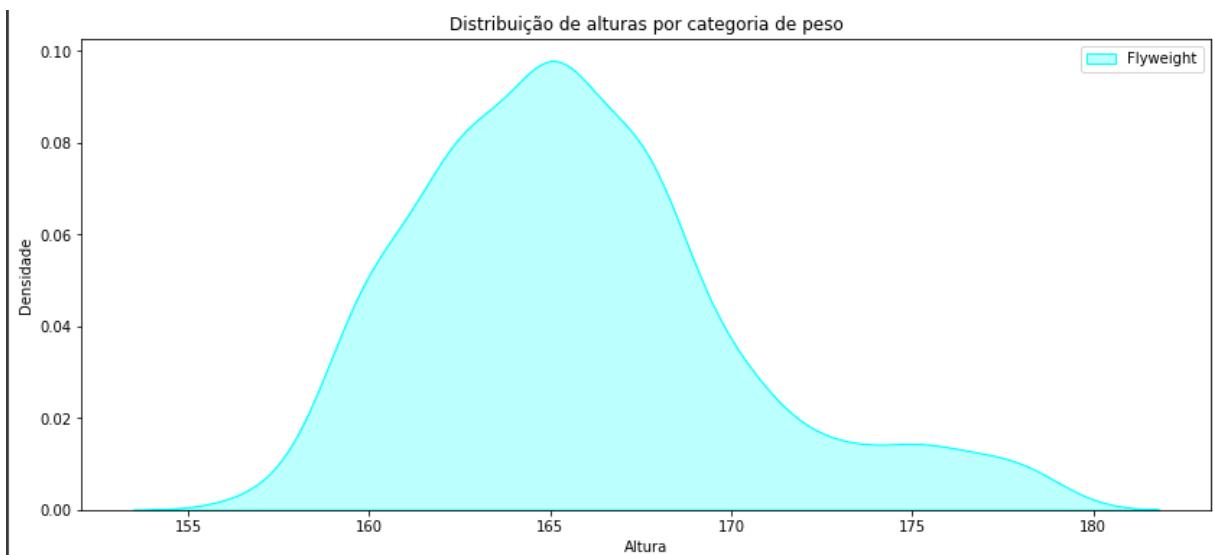
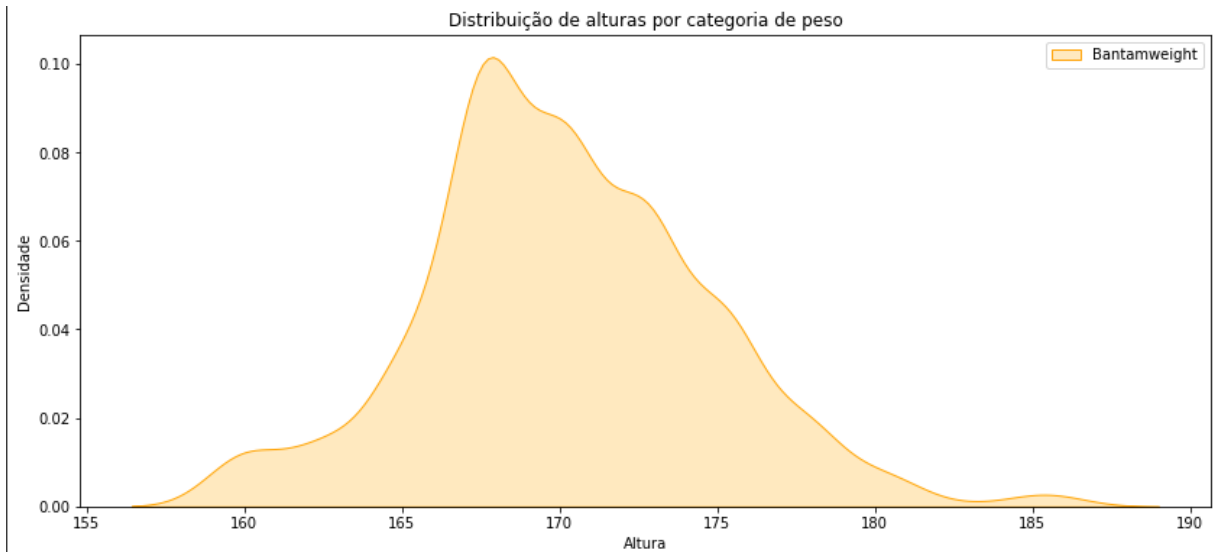


Os próximos gráficos mostram a distribuição de altura dos atletas organizados por categoria. Como é de se esperar a categoria com lutadores mais altos é a mais pesada









## 7. Links

Link para o código python no google colab notebooks: [clique aqui](#)

Link para o repositório no github com os códigos: [clique aqui](#)

Link para a apresentação: [clique aqui](#)

## REFERÊNCIAS

PESSANHA, Cíntia. Random Forest: como funciona um dos algoritmos mais populares de ML. **Medium**, **2019**. Disponível em: <https://medium.com/cinhiabpessanha/random-forest-como-funciona-um-dos-algoritmos-mais-populares-de-ml-cc1b8a58b3b4>. Acesso em: 26 de fevereiro de 2021.

RANA, Kartikeya. Become a Pokémon Master with Machine Learning. **Towards Data Science**, **2020**. Disponível em: <https://towardsdatascience.com/become-a-pok%C3%A9mon-master-with-machine-learning-f61686542ef1>. Acesso em: 10 de fevereiro de 2021.

VALKOV, Venelin. **Hacker's Guide to Neural Networks in Javascript**. Kindle, Amazon, 2019-2020.