

MGT 6203 Group Project Proposal

TEAM INFORMATION (1 point)

Team #: 7

Team Members: name + gtid + [Insert background information: Name, professional background, education background, previous analytics related projects you have worked on]

1. Andrew Berkowitz [aberkowitz8]: Undergrad in business currently working as a senior data analyst.
2. Alejandro Forero Eng [aeng61]: Undergrad in Electrical Engineering, master's in electronic and computers, and MBA, currently working as Director of Technology.
3. Juliette Noelle Wong [jwong88]: Undergrad in math & statistics, past projects include analyzing the effect of age on defensive ability at different positions for a Major League Baseball team. Currently working as a software quality engineer/data analyst.
4. Nicholas P Wilson [nwilson86]: Undergrad in Economics, currently working as a forecasting analyst in the energy industry.
5. Samuel Wade Wang [swang3068]: Was a graduate student majoring in pure math. During years in OMSA, worked on a [flu spread simulation project](#) in Simulation course; currently also working on course projects from DVA, CDA, DL this semester.

OBJECTIVE/PROBLEM (5 points)

Project Title: UFC Fight Prediction

Background Information on chosen project topic:

The Ultimate Fighting Championship (UFC) is the most well-known organization in the growing sport of mixed martial arts. As the sport and organization grow in popularity, there will be more people betting on fight outcomes, which increases the amount of people watching the events. We would like to identify key features that predict the winner of the fight. Additionally, we would like to explore if there are fighters and fight history that affect how many PPVs (Pay Per View) an event sells and create a model to predict the number.

Problem Statement (clear and concise statement explaining purpose of your analysis and investigation):

What are the major factors in determining which fighter wins, and how can we use them to predict the results of future bouts?

State your Primary Research Question (RQ):

What factors are the strongest determinants of which opponent wins a UFC match?

1. Factor analysis on UFC fight data for PPV (Pay Per View Buys).
2. Given a fighter's past fight statistics, can we predict who will win the match?
3. What would be the best strategy to maximize the odds of winning a fight?

1. The UFC could benefit from a model showing largely objective measures determining fight outcomes. This affects fighter income, betting odds, and perceived fairness of the brand/league.
2. Predictive portions would allow UFC and other entities to better understand what drives PPV sales.
3. Sports analytics in general are also influential in sports betting industries.

- Primary Datasource: UFC fight data from 1993 to 2023
<https://www.kaggle.com/datasets/akshaysinghim/ufc-fight-data-1993-to-2023?resource=download>
- Secondary Datasource: UFC PPV Sales
<https://www.kaggle.com/datasets/daverosenman/ufc-ppv-sales>

		G										H										I										J										K										L										M										N										O										P										Q										R										S										T										U										V										W										X										Y										Z										AA										AB										AC										AD										AE										AF										AG										AH										AI										AJ										AK										AL										AM										AN										AO										AP										AQ										AR										AS										AT										AU										AV										AW										AX										AY										AZ										BA										BB										BC										BD										BE										BF										BG										BH										BI										BJ										BK										BL										BM										BN										BO										BP										BQ										BR										BS										BT										BU										BV										BW										BX										BY										BZ										CA										CB										CC										CD										CE										CF										CG										CH										CI										CJ										CK										CL										CM										CN										CO										CP										CQ										CR										CS										CT										CU										CV										CW										CX										CY										CZ										DA										DB										DC										DD										DE										DF										DG										DH										DI										DJ										DK										DL										DM										DN										DO										DP										DQ										DR										DS										DT										DU										DV										DW										DX										DY										DZ										EA										EB										EC										ED										EE										EF										EG										EH										EI										EJ										EK										EL										EM										EN										EO										EP										EQ										ER										ES										ET										EU										EV										EW										EX										EY										EZ										FA										FB										FC										FD										FE										FF										FG										FH										FI										FJ										FK										FL										FM										FN										FO										FP										FQ										FR										FS										FT										FU										FV										FW										FX										FY										FZ										GA										GB										GC										GD										GE										GF										GG										GH										GI										GJ										GK										GL										GM										GN										GO										GP										GQ										GR										GS										GT										GU										GV										GW										GX										GY										GZ										HA										HB										HC										HD										HE										HF										HG										HH										HI										HJ										HK										HL										HM										HN										HO										HP										HQ										HR										HS										HT										HU										HV										HW										HX										HY										HZ										IA										IB										IC										ID										IE										IF										IG										IH										II										IJ										IK										IL										IM										IN										IO										IP										IQ										IR										IS										IT										IU										IV										IW										IX										IY										IZ										JA										JB										JC										JD										JE										JF										JG										JH										JI										JJ										JK										JL										JM										JN										JO										JP										JQ										JR										JS										JT										JU										JV										JW										JX										JY										JZ										KA										KB										KC										KD										KE										KF										KG										KH										KI										KJ										KK										KL										KM										KN										KO										KP										KQ										KR										KS										KT										KU										KV										KW										KX										KY										KZ										LA										LB										LC										LD										LE										LF										LG										LH										LI										LJ										LK										LL										LM										LN										LO										LP										LQ										LR										LS										LT										LU										LV										LW										LX										LY										LZ										MA										MB										MC										MD										ME										MF										MG										MH										MI										MJ										MK										ML										MM										MN										MO										MP										MQ										MR										MS										MT										MU										MV										MW										MX										MY										MZ										NA										NB										NC										ND										NE										NF										NG										NH										NI									
--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	---	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--

	A	B	C	D	E	F	G
1	Year	Month	Day	UFC_Num	Opponent	Opponent	PPV
2	2001	9	28	33	Ortiz	Matyushe	75000
3	2001	11	2	34	Couture	Rizzo	65000
4	2002	1	11	35	Pulver	Penn	35000
5	2002	3	22	36	Couture	Barnett	55000
6	2002	5	10	37	Lindland	Bustaman	50000
7	2002	7	13	38	Hughes	Newton	45000
8	2002	9	27	39	Couture	Rodriguez	45000
9	2002	11	22	40	Ortiz	Shamrock	100000
10	2003	2	28	41	Rodriguez	Sylvia	60000
11	2003	4	25	42	Hughes	Sherk	35000
12	2003	6	6	43	Couture	Liddell	49000
13	2003	9	26	44	Couture	Ortiz	94000
14	2003	11	21	45	Hughes	Trigg	40000
15	2004	1	31	46	Couture	Belfort	80000
16	2004	4	2	47	Ortiz	Liddell	105000
17	2004	6	19	48	Shamrock	Kimo	110000
18	2004	8	21	49	Couture	Belfort	80000
19	2004	10	22	50	Ortiz	Cote	40000
20	2005	2	5	51	Ortiz	Belfort	105000

Key Variables: (which ones will be considered independent and dependent? Are you going to create new variables? What variables do you hypothesize beforehand to be most important?)

- Dependent
 - Winner
 - PPV
- Independent
 - Knockdowns
 - Significant strikes
 - Total strikes
 - Takedowns
 - Submission attempts
 - Control time
 - Type of strike (head, body, leg)
 - Striking range (distance, clinch, ground)
 - Winning method
 - Fight type (title bout, bout, etc.)
- Created Variables
 - Record
 - Interaction Terms
 - Dummy Variables

For predicting the winner, given the fight goes to decision, we believe the most significant variables will be knockdowns, total strikes, takedowns, and control time. For predicting PPV sales, we believe that a fighter's previous fight outcomes will be most important. If they have been ending fights or on a win streak, they may be a bigger draw.

APPROACH/METHODOLOGY (8 points)

Planned Approach (In paragraph(s), describe the approach you will take and what are the models you will try to use? Mention any data transformations that would need to happen. How do you plan to compare your models? How do you plan to train and optimize your model hyper-parameters?))

We will approach the problem in a traditional and methodical manner.

Integration of dataset: Starting with a merge of both datasets by date, ensuring that each fight's statistics are paired with its corresponding PPV sales figure. This will enable us to answer our additional PPV research question.

Cleaning and transformation: Data cleaning steps will include splitting some string-type predictors to quantitative values (e.g. R_SIG_STR., R_TD), removing collinear columns (e.g. R_SIG_STR_pct, R_TD_pct), decide how to handle missing values (e.g. case of time 0:00 for both fighters). Transformation will involve normalizing numerical variables (e.g., strikes, takedowns) to ensure they are on a similar scale, encoding categorical variables (e.g., winning method, fight type) using one-hot encoding, and aggregating fighter statistics to create comprehensive profiles.

Features Engineering: Considering the number of features in our datasets, and how they grow when we transform our data, we may need to apply some reduction dimensionality techniques such as PCA or recursive analysis elimination.

Model Selection: For predicting fight outcomes, algorithms like logistic regression, decision trees and random forests could be considered. For PPV sales predictions, linear regression models and random forests considering the continuous nature of the target variable. AIC could be applied to compare and select models.

Training and Validation: Models will be trained using a split of the dataset into training and validation sets. Cross-validation, specifically k-fold cross-validation, will be employed to assess model robustness and generalizability, ensuring that performance is consistent across different subsets of the data.

Anticipated Conclusions/Hypothesis (what results do you expect, how will you approach lead you to determining the final conclusion of your analysis) Note: At the end of the project, you do not have to be correct or have acceptable accuracy, the purpose is to walk us through an

analysis that gives the reader insight into the conclusion regarding your objective/problem statement

Prior to our analysis, we expect to be able to identify several features that are statistically significant in predicting the winner of a fight. Features that we think will have the most significance are knock downs, total strikes, control time, and take downs. We then expect to be able to predict the winner of fights with reasonable accuracy. For our secondary question of PPV sales, we expect fighters on a win streak, fighters with a record of knockouts or submissions, and whether the event is a milestone (UFC 100, etc.) are important in PPV sales.

What business decisions will be impacted by the results of your analysis? What could be some benefits?

Through our analysis, we will better understand what measures are significant in determining fight outcomes. Given those determinants, teams would be able to augment their strategies. This can also help individual bettors have an advantage to win and give the sports betting industry a tool to help calculate betting odds. Uncovering what drives higher PPV sales will further improve the UFC's ability to generate revenue from events.

PROJECT TIMELINE/PLANNING (2 points)

Project Timeline/Mention key dates you hope to achieve certain milestones by:

Some possible items for data analysis:

1. Complete data cleaning: 1 week
2. Exploratory data analysis: 3 weeks
3. Fitting models to the data, record findings: 3 weeks
4. Turning findings into report: 1 week

Some possible items for documents

1. Proposal (2/25)
2. Progress Report (3/17)
3. Final Report (4/21)

Appendix (any preliminary figures or charts that you would like to include):