

Poročilo za prvo seminarsko nalogo
Odločitveni model za izbiro najbolj vrednega igralca v NBA

1. Opis problema

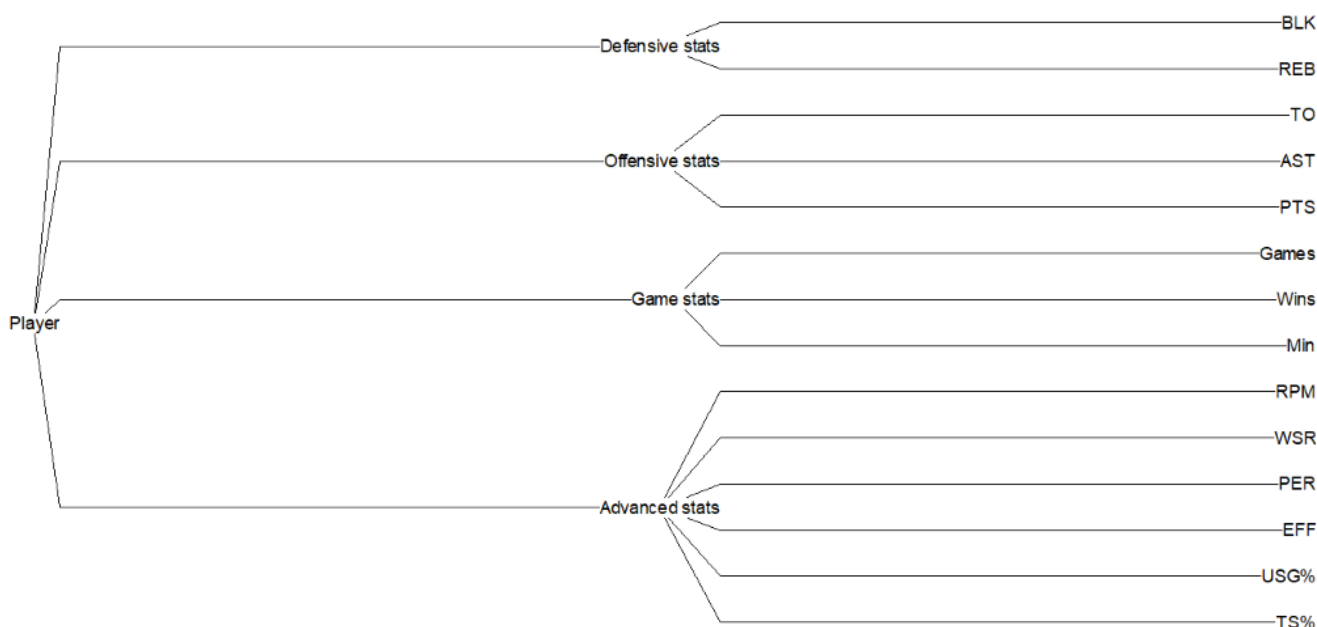
Za odločitveni problem sem pri tej seminarski nalogi določil izbiro najboljšega košarkarja v ameriški NBA ligi. Za to temo sem se odločil, ker sem osebno navdušenec NBA košarke, menim pa da v zadnjih letih popularnost le-te v Sloveniji narašča, najbolj zaradi uspešnosti dveh slovenskih košarkarjev Gorana Dragiča ter Luke Dončića.

Najbolj vreden igralec (ang. MVP) je v ligi NBA najbolj prestižna osebna nagrada, ki jo lahko igralec prejme. V splošnem se nagrada podeljuje po koncu redne sezone med končnico, zmagovalec je pa izbran po procesu volitev, kjer člani medijskih hiš ter športni novinarji v USA in Kanadi podajo svoje prve štiri izbire. Med temi štirimi, igralec prejme 10 točk za prvo mesto, 7 za drugo, 5 za tretje ter 1 za četrto mesto. Te glasovi so subjektivni in po navadi kontroverzni ter v veliki večini sprožijo zagrete pogovore med oboževalci košarke.

Skozi leta sta za izbiro tega igralca dominirali kategoriji "Najboljši igralec v najboljši ekipi" ter "Igralec, ki v sezoni doseže zgodovinske statistike". Moj cilj v tej seminarski je da z uporabo statistike košarkarjev v sezoni 2019/2020 določim najbolj vrednega izmed vseh.

2. Razlaga kriterijev in izpeljanih vozlišč v drevesu

Odločitveni model vsebuje 14 osnovnih kriterijev ter 4 izpeljana vozlišča.



Slika 1: Drevo kriterijev

Kriteriji v izpeljanem vozlišču »Defensive stats« opisujejo igralčevo statistiko v obrambi. »BLK« označuje blokirane mete, »REB« pa uspešne skoke za žogo po metu igralca pred svojim ali nasprotnem košom .

»Offensive stats« vsebuje osnovno statistiko o uspešnosti košarkarja v napadu. »TO« ali »Turnovers« oz. preobrati nam pove kolikokrat je igralec v napadu izgubil žogo, bodisi s slabo podajo bodisi zaradi kraje nasprotnika. Koristnost tega kriterija je inverzna, torej je manj preobratov bolje je za igralca. Kriterija »AST« določuje igralčeve podaje, ki so bile direktno spremenjene v točke soigralca »PTS« pa opisuje igralčeve točke.

Vozlišče »Game Stats« vsebuje kriterije, ki nam dajo informacijo o igralčevi sezoni. »Min« nam pove koliko minut ter sekund je košarkar igral na igrišču. Ta kriterij, tako kot vsi do sedaj naštetih v vozliščih »Offensive stats« ter »Defensive Stats« so normalizirani na 48 minut torej trajanje ene igre. Kriterija »Games« ter »Wins« enostavno opisujeta koliko iger je košarkar igral ter koliko jih je zmagal.

Vozlišče »Advanced stats« je obenem najbolj pomembno ter tudi najbolj kompleksno. To vozlišče vsebuje kriterije, ki opisujejo napredno statistiko košarkarja, ki je tudi izračunana iz drugih bolj osnovnih podatkov. Ker te kriteriji pokrivajo velik del košarkarjeve igre, je tudi to vozlišče v modelu najbolj pomembno z največjo utežjo.

Kriterij »RPM« ali realistični plus minus ocenjuje igralčev vpliv na parketu v njegovi ekipi. Normaliziran je na absolutno razliko točk na 100 napadalnih in obrambnih posesti, opoštuje pa tako košarkarjeve soigralce, nasprotnike ter druge faktorje.

»WSR« oz. Razmerje zmagovalnega deleža opisuje število zmag, ki jih je košarkar sam doprinesel ekipi, normaliziran na 48 minut. Povprečje v ligi je približno 0,1 oz. 10%, kar pomeni da povprečen igralec v ligi doprinese k 10% zmag v ekipi.

»PER« ali faktor košarkarjeve učinkovitosti ocenjuje igralčevo produktivnost na minuto. Ta statistika je izpeljana iz veliko osnovnih pozitivnih kot so podaje, skoki, bloki, kraje, meti za 3 točk, meti z polja, prosti meti ter tudi negativnih npr. zgrešeni meti, preobrati ter osebni prekrški. Formula doda pozitivne podatke ter odšteje negativne z uporabo statističnega točkovnega sistema. Povprečje v ligi je 15,0.

»USG%«, igralčev delež uporabe izračuna kakšen delež ekipnih napadalnih posesti je bil košarkar vključen, pod pogojem, da se posest zaključi z enim izmed treh rezultatov: met iz polja, prosti met ali preobrat. Povprečen igralec bo imel ta delež enak 20 odstotkov, kar je smiselno glede na to da je na parketu na enkrat 5 igralcev ene ekipe.

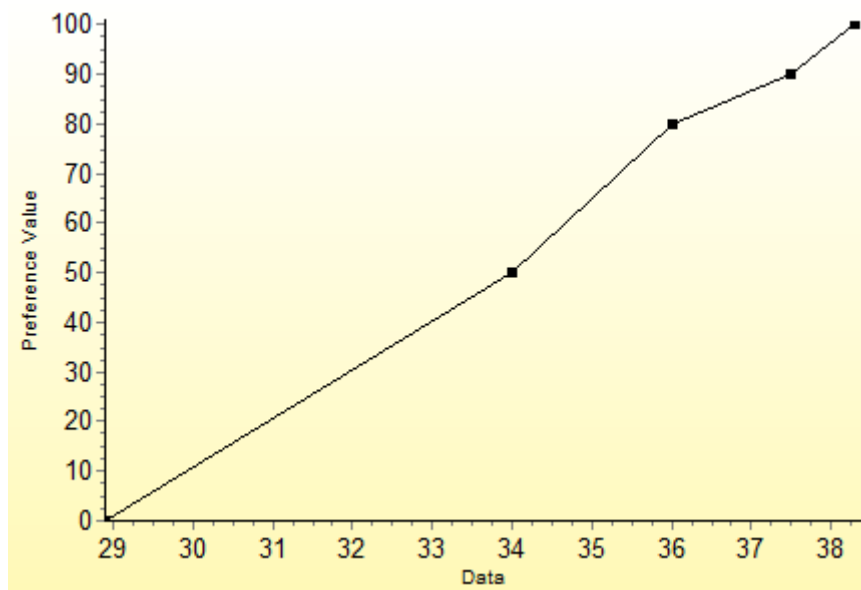
»TS%« je zadnji kriterij modela. »True shooting percent« ali realistični odstotek meta računa košarkarjevo učinkovitost pri metu žoge. Vključuje mete iz polja, mete za 3 točke ter proste mete.

3. Razlaga uteži in funkcij koristnosti

Za večino atributov sem za izračun koristnosti uporabil navadno relativno funkcijo koristnosti.

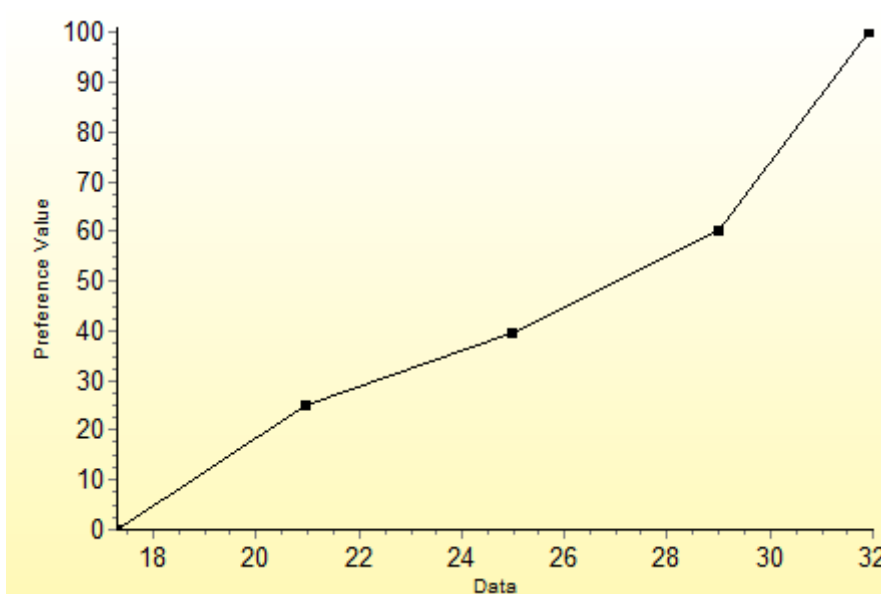
Pri kriterijih »PER« ter »USG%« sem uporabil odsekoma linearno premico, saj sem hotel nagraditi košarkarje z visokimi vrednostmi, ki so dosegli zgodovinske statistike.

Pri kriteriju delež uporabe, so bili kar trije košarkarji (James Harden, Giannis Antetokounmpo, Luka Dončić) nad vrednostjo 36.00, kar jih uvršča med top 10 vodilnih deležev uporabe v sezoni odkar se vodi ta statistika (1977-78).



Slika 2: Funkcija deleža uporabe (USG%)

Pri faktorju učinkovitosti sem še posebej hotel izpostaviti zmagovalca med vsemi, saj je Giannis Antetokounmpo po »PER« kriteriju imel najboljšo sezono v zgodovini NBA, kjer se je z vrednostjo 31.86 prebil kar na prvo mesto in tako prehitel druge legende NBA kot so Wilt Chamberlain, Michael Jordan ter ostali. PER nad 31.2 bi pomenil mesto v top 10 vseh časov, James Harden pa se je na drugem mestu z vrednostjo 29.14 uvrstil na 41. mesto vseh časov.



Slika 3: Funkcija faktorja učinkovitosti (PER)

Med vsemi izpeljanimi vozlišči mi je najbolj pomemben »Advanced stats« za njim pa sledita »Offensive stats« ter »Game stats«. »Defensive stats« sem postavil na zadnje mesto, saj sem se hotel osredotočiti na napadalnost košarkarjev bolj kot obrambo, ker se na le-to osredotoča nekoliko drugačna a še vseeno spoštovana nagrada (DPOY), ki določi obrambnega košarkarja sezone.

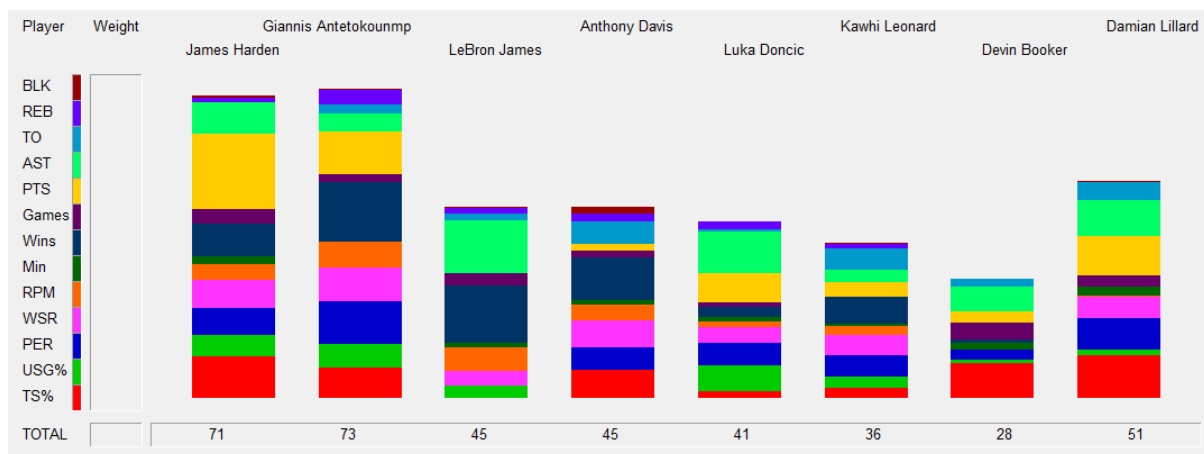
V »Advanced stats« sta mi najbolj pomembna odstotek meta (TS%) ter faktor košarkarjeve učinkovitosti (PER). Odstotek meta je pomemben ker opisuje dejansko spretnost igralčevega meta, ki je zelo pomembna vendar ni bolj pomembna kot faktor učinkovitosti, ker upošteva tudi proste mete. Le-te si nekateri igralci zelo spretno znajo priboriti, včasih tudi s prevaro sodnikov, kar je po mojem mnenju pomankljivost te statistike. Faktor učinkovitosti mi je zelo pomemben saj z eno vrednostjo splošno opisuje igralčevo igro, zaradi upoštevanja tudi negativnih košarkarjevih učinkov.

Naslednji kriterij po pomembnosti je razmerje zmagovalnega deleža igralca (WSR). Ta podatek nam jasno pove koliko je košarkar doprinesel k zmagi ekipe in v tem športu so zmage tisto kar je najbolj pomembno. Ta kriterij sem utežil tako ker sem hotel nekako kaznovati igralce, ki so zelo dobri v metu žoge (TS%) ampak morda ne doprinesejo precej k ekipi, kar je po mojem mnenju bolj pomembno v tem športu.

Zadnja napredna kriterija sta realistični plus minus (RPM) ter delež uporabe (USG%). Čeprav ne za veliko (5%) menim, da sta manj pomembna od drugih naprednih kriterijev saj imata oba svoje pomankljivosti. Kot prej omenjeno sem hotel v tem modelu dati poudarek na napadalnih spretnostih košarkarjev, kar pa ne drži popolnoma za RPM. Le-ta enakomerno upošteva napadalni »RPM« ter obrambni »RPM« in zato sem ga utežil z manjšo vrednostjo.

Igralčev delež uporabe (USG%) nam da približen opis o tem kako ekipa uporablja košarkarja. Pomankljivost te statistike je ta, da v svojem izračunu ne upošteva košarkarjevih asistenc, zaradi česar sem ga utežil kot najmanj vrednega.













4. Najboljša varianta in primerjava alternativ



Slika 4: Primerjava variant

Najboljši igralec v tem modelu je Giannis Antetokounmpo, kateremu tesno sledi James Harden. Ostali so si med sabo precej podobni in se gibljejo okoli ocene 45, dve izjemi sta Kawhi Leonard ter Devin Booker, ki je precej oddaljen od drugih na zadnjem mestu.

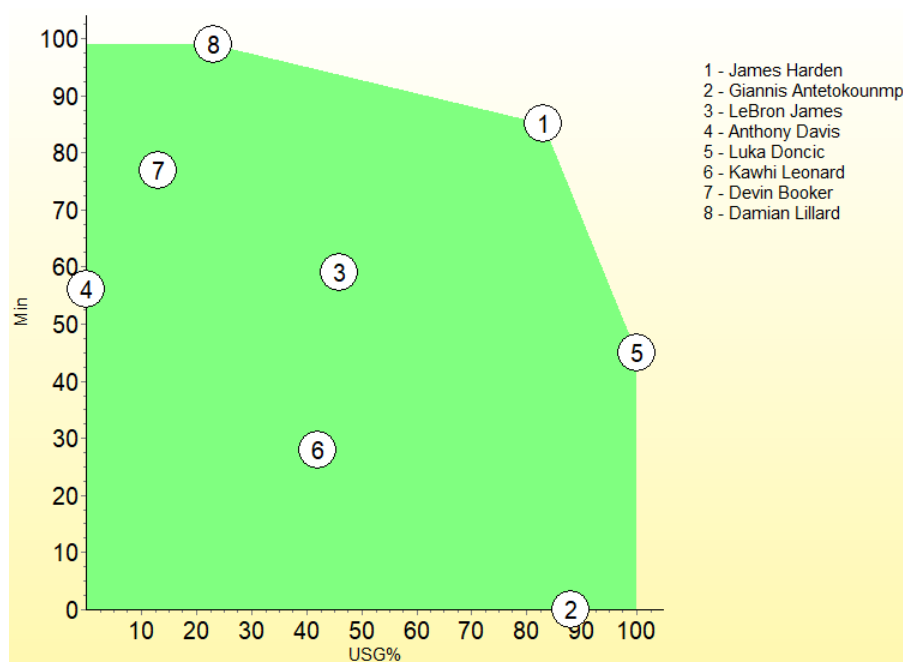
Diagram nam prikazuje kje sta Giannis in James prejela največji delež svojih točk. Vidimo, da je k velikemu delu James-ove ocene prispeval atribut »PTS«, kar ni presenetljivo glede na to, da je Harden v zadnjih dveh letih povprečno na tekmo dosegel kar 35 točk. Giannis-u so največ prispevale zmage, katerih je imel 8 več kot James.

	Model Order	Cum Wt	Diff	Wtd Diff	Sum	
Game stats	Wins	14,0	44	6,2	6,2	
Advanced stats	PER	10,0	38	3,8	10,1	
Defensive stats	REB	3,3	74	2,5	12,5	
Advanced stats	RPM	6,0	36	2,2	14,7	
Offensive stats	TO	5,3	40	2,1	16,8	
Advanced stats	WSR	8,0	18	1,5	18,3	
Advanced stats	USG%	6,0	5	0,3	18,6	
Defensive stats	BLK	1,7	5	0,1	18,7	
Game stats	Games	4,0	-38	-1,5	17,1	
Game stats	Min	2,0	-86	-1,7	15,4	
Advanced stats	TS%	10,0	-26	-2,6	12,8	
Offensive stats	AST	12,3	-27	-3,3	9,5	
Offensive stats	PTS	17,5	-43	-7,6	1,9	
		100.0		1.9		

Slika 5: Primerjava variant

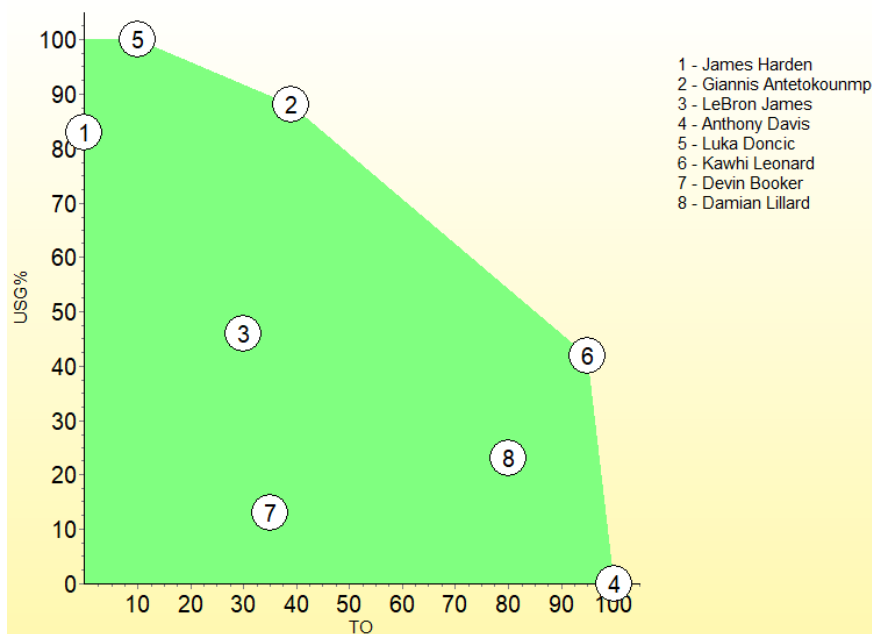
Primerjava med Giannis-om in njegovo najboljšo alternativo James-om nam pokaže kje se igralca najbolj razlikujeta. Kot prej omenjeno sta atributa »PTS« ter »Wins« daleč najbolj prispevala k vsakemu posamezniku. Vidimo tudi da James bolje asistira soigralcem kot pričakovano pa ima tudi zelo dober met žog (TS%). Giannis je veliko bolj učinkovit v svoji igri prav tako pa ima precej več točk nad svojo zamenjavo (RPM). Prav tako je Giannis precej manj nagnjen k izgubi žoge kar poudari njegovo konsistenco.

5. Analiza



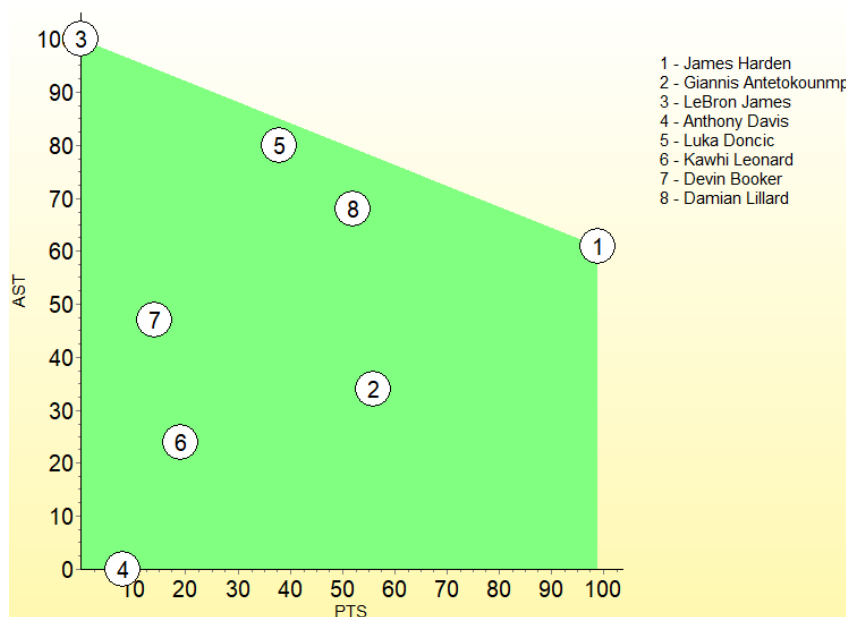
Slika 6: Uporabnost igralca in igrane minute

Zgornji graf nam prikazuje igralčevo uporabnost ter njegove igrane minute. Vidimo da je Giannis daleč zadnji v igranih minutah na igro ampak še vseeno sodeluje v velikem deležu ekipnih posesti. To potrди tudi njegov realistični plus minus »RPM« pri vrednosti 10.3, kateri nam pomaga dobiti idejo o njegovi pomembnosti v ekipi.



Slika 7: Preobratni igralca ter njegova uporabnost

Na grafu vidimo kako pogosto igralci izgubljajo žogo v napadu. Anthony Davis je zelo dober v tej statistiki saj ne povzroča veliko preobratov ampak je prav tako najmanj uporabljen v ekipi. Na drugem koncu je Luka Dončić zelo uporaben v svoji ekipi ampak velikokrat izgubi žogo, kar pa mu redki očitajo glede na to da je zelo mlad.



Slika 8: Igralčeve točke in asistence

Ta graf opisuje kako nagnjeni so igralci k metu ali podaji žoge. James Harden, ki je večš strelec je kot pričakovano precej nagnjen k metu žoge a še vseeno komplementira svojo vlogo z relativno visokim številom asistenc na igro. LeBron James na drugi strani v svoji 17. sezoni morda ni več tako neusmiljen strelec kot je bil včasih v vrhuncu svoje kariere a je vseeno zelo koristen v svoji ekipi z kar 10 podajami na igro.

Igralci ki so blizu zgornje črte so zelo učinkoviti organizatorji igre in velikokrat dosežejo dvojne dvojčke (10+ točk, 10+ podaj) ali celo trojne dvojčke (10+skokov), ki so v košarki precej redki.

6. Kaj-če analiza

V tej analizi me je zanimalo koliko več tekem bi moral James zmagati, da bi prehitel Giannis-a v modelu. Za ta kriterij sem se odločil, ker ga lahko neposredno povežemo z kvaliteto Houston Rockets ekipe, kjer James Harden igra. Moje vprašanje je torej: »Kaj če bi bila ekipa James-a boljša?«. Takšen scenarij ima lahko morebitne druge posledice kot znižanje James-ovih drugih statistik npr. uporabnost v ekipi, realistični plus minus a sem pri tej analizi predpostavljal, da le-te ostanejo enake.

Trenutno ima Giannis pri 51 zmagah 8 več zmag kot James. Pri 48 zmagah se James izenači z Giannis-om torej potrebuje samo 5 zmag. Pri naslednji zmagi James prevzame vodstvo.