

Tartu Ülikool
Loodus- ja täppisteaduste valdkond
Matemaatika ja statistika instituut

NBA mängijad

Projekt

Autorid: Raigo Luhaorg
Joonas Anton Jürgenstein

Tartu 2020

Sissejuhatus

Töö eesmärgiks on analüüsida andmestikku, mis sisaldab andmeid korvpalliliigas NBA aastatel 1996-2020 mänginud mängijate kohta. Autorid püstitasid 2 põhilist küsimust:

1. Kas mängija pikkus mõjutab tema keskmiselt visatud punktide arvu mängus ja kuidas see on muutunud aja jooksul?
2. Millisest ülikoolist on tulnud kõige rohkem NBA mängijaid ning milline on nende keskmine +- reiting?

Lisaks uurisime, kuidas on NBAs mängijate vanus hooaegade lõikes muutunud ja ka mõningaid muid näitajaid, mis meile huvi pakkusid. Andmestik on saadud veebileheküljelt [kaggle.com](https://www.kaggle.com).

Andmestiku kirjeldus

Uuritud objektid

Uuritavateks objektideks on korvpalliliigas NBA mänginud korvpallurid aastatel 1996-2020, kelle kohta on andmestikus välja toodud nii tema füüsilised kui ka statistilised näitajad erinevate hooegade jooksul. Kokku on andmestikus 11145 rida, millest unikaalsed ehk erinevaid mängijaid on kokku 2235.

Uuritud tunnused

Kokku on andmestikus 22 erinevat tunnust: X - index

player_name - mängija nimi

team_abbreviation - tiimi nime lühend

age - vanus

player_height - mängija pikkus

player_weight - mängija kaal

college - ülikool, kus mängija enne NBAse siirdumist mängis

country - päritolu riik

draft_year - mis aastal mängija liigasse valiti

draft_round - mitmendas ringis mängija NBAse valiti

draft_number - mitmenda valikuna mängija NBAse valiti

gp - mängitud mängu

pts - keskmiselt visatud punkte

reb - keskmiselt võetud lauapalle

ast - keskmiselt jagatud resultatiivseid sööte

net_rating - väljakul oldud aja jooksul meeskonna skoor võrreldes vastasega

oreb_pct - osakaal väljakul oldud aja jooksul teenitud ründelauapallidest

dreb_pct - osakaal väljakul oldud aja jooksul teenitud kaitselauapallidest

usg_pct - osakaal, palju mängijat kasutati tema väljakul oldud aja jooksul

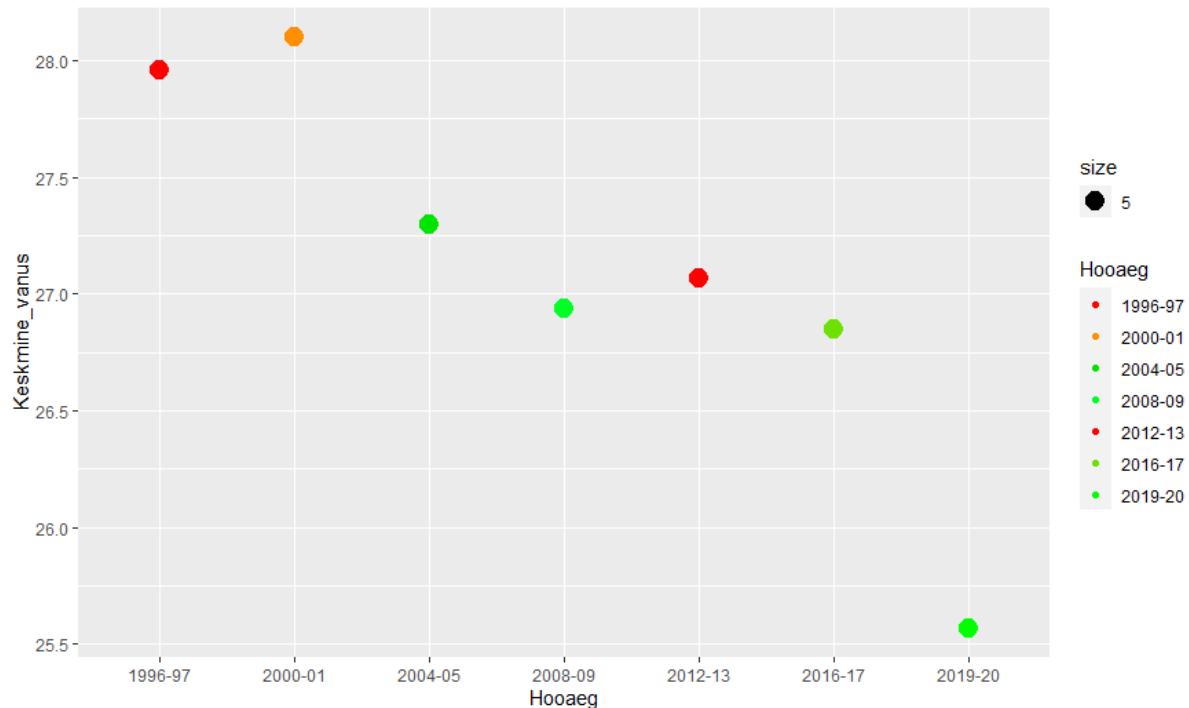
ts_pct - visketabavuse efektiivsus

ast_pct - osakaal, palju mängija jagas resultatiivseid sööte enda mängijale, kui ta oli väljakul

season - mängitud hooaeg

Materjali kirjeldav analüüs

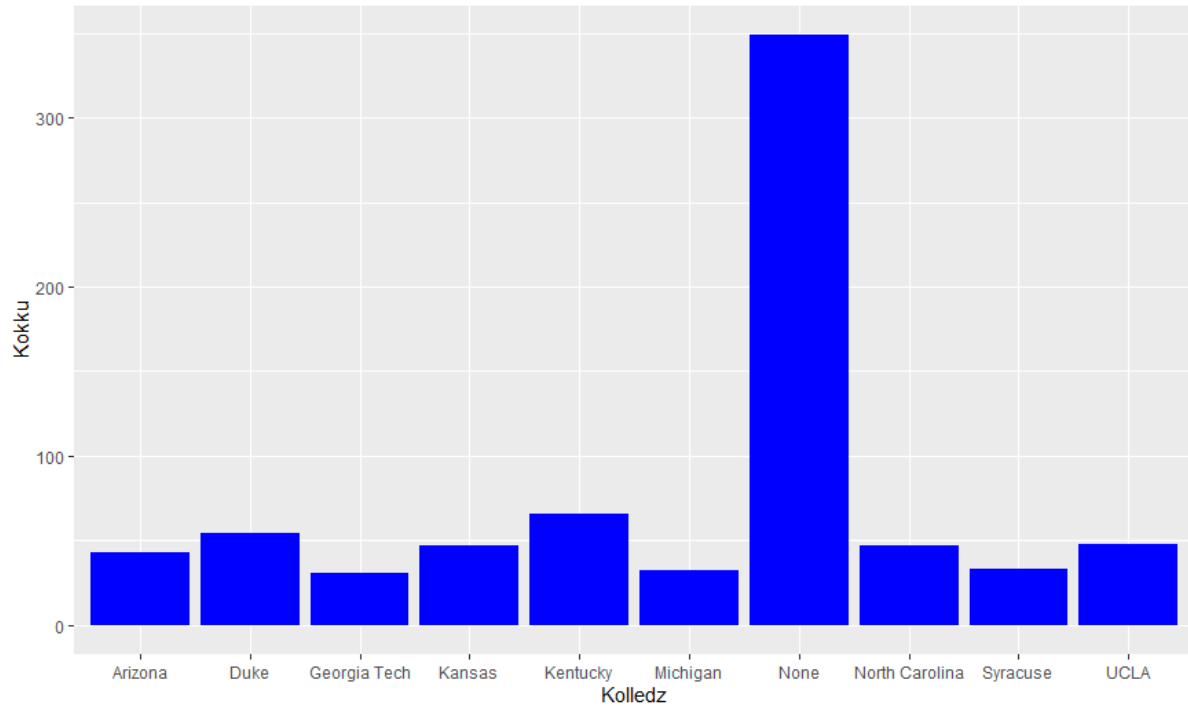
Andmestikus olevad andmed ja statistika on kogutud alates 1996-1997 hooajast kuni käesoleva hooajani. Samas peab tõdema, et kõikide andmestikus olevate mängijate kohta on olemas vaid tema mõne hooaja statistilised näitajad, samas kui mõne teise mängija kohta on olemas seal terve tema NBA karjääri kõikide hooegade statistika. Vaatleme näiteks NBAs mängijate keskmist vanust erinevatel hooaegadel. Võtame näitajad iga 4 hooaja kaupa.



Joonis 1. Mängijate keskmine vanus hooegade lõikes

Nagu tabelist näha, siis viimase 25 hooaja jooksul on NBA mängijate keskmine vanus hooegade jooksul vähenenud ligikaudu 2 aasta võrra ehk liigasse on rohkem juurde tulnud nooremaid mängijaid. Kõige kõrgema vanusega, 28.1 aastat, oli hooajal 2000-01. Kõige madalam keskmine vanus alates hooajast 1996-97 on olnud 25.57 aastat ning selline keskmine vanus oli läinud NBA hooajal ehk hooajal 2019-20.

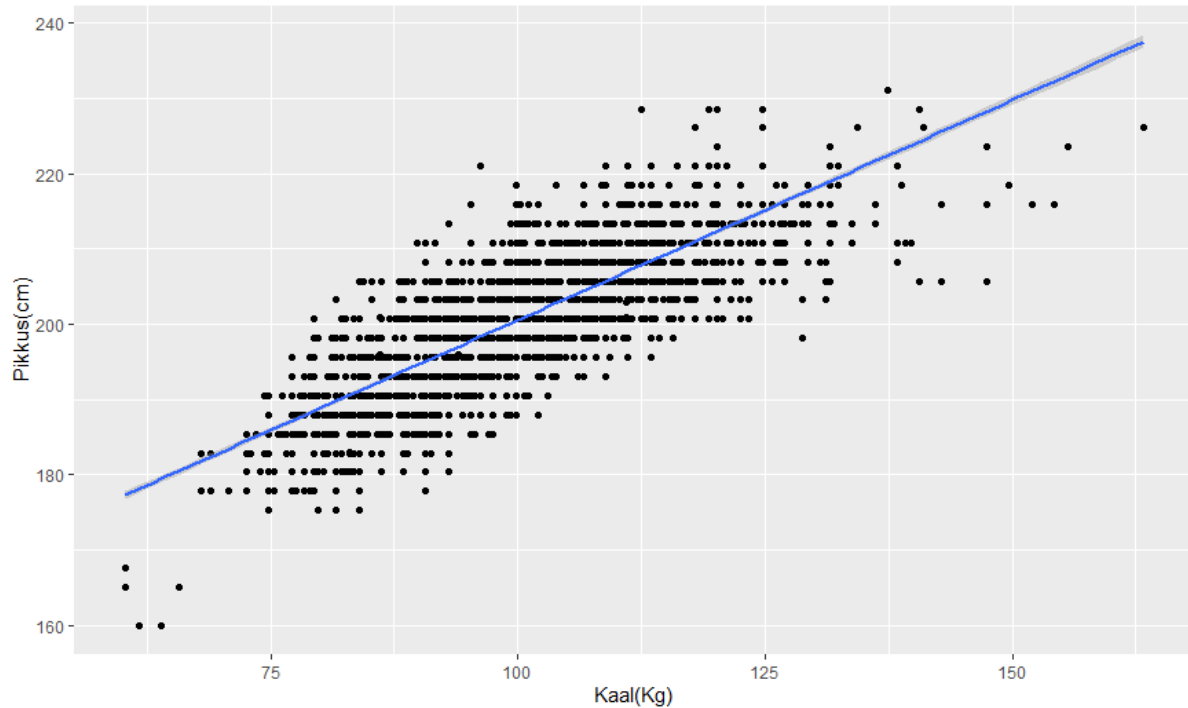
Vaatleme nüüd ka, millist NBA kolledžist on korvpalliliigasse NBA läinud kõige rohkem mängumehi. Lihtsuse huvides valisime 10 kõige populaarsemat ülikooli, kus mängijad on liigasse läinud.



Joonis 2. Millisest kolledžist on kõige rohkem mängijaid tulnud

Histogrammi põhjal näeme aga hoopis, et enamus mängijad pole liigasse siirdunud Ameerika ülikoolist, vaid hoopis mujalt. Üldjuhul on sellisel juhul tegemist eurooplasega või on mängija läinud NBAse otse keskkoolist. Selliseid mängijaid on kokku 349. Kui aga vaadata Ameerika ülikoolist NBAse siirdunud mängumehi, siis enim on käinud Kentucky ülikoolis, kus on kokku käinud 66 NBA mängijat.

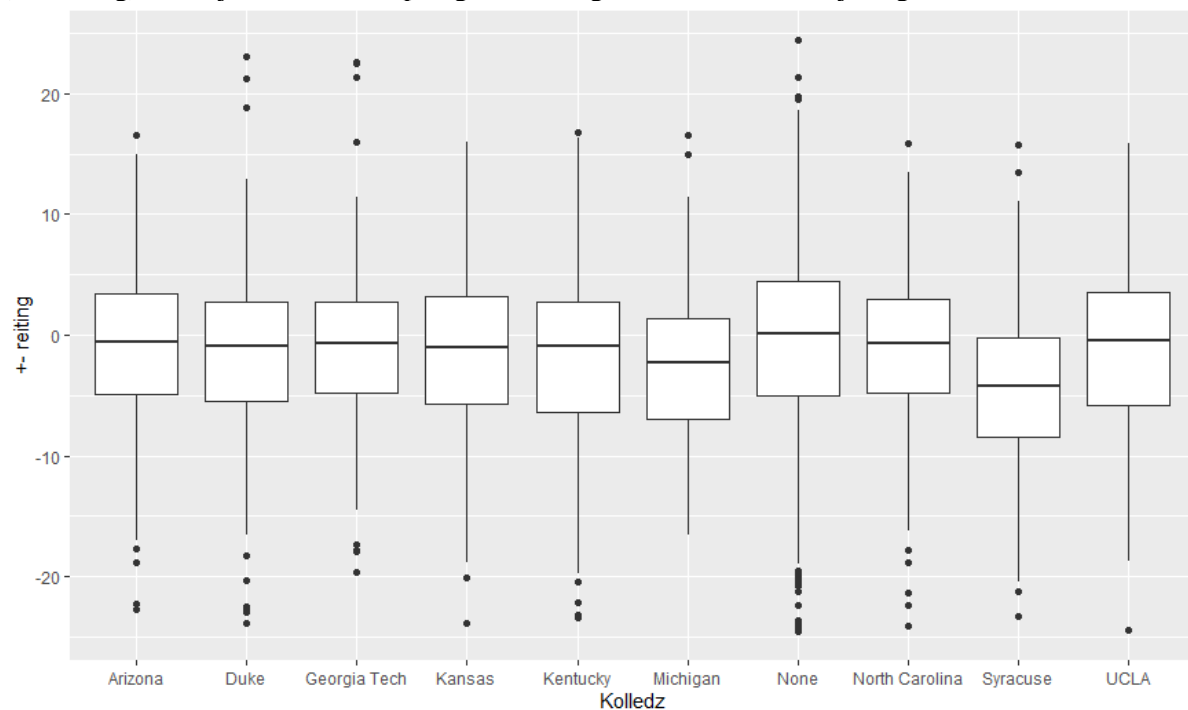
Uurime nüüd, milline on mängijate kaalu ja pikkuse suhe. Joonisel peab ka arvestama, et mõne mängija kohta mitu täppi, kuna hooegade lõikes on mõnede mängijate kaal või pikkus muutunud.



Joonis 3. Mängijate pikkuse ja kaalu suhe ning selle regresioonisirge

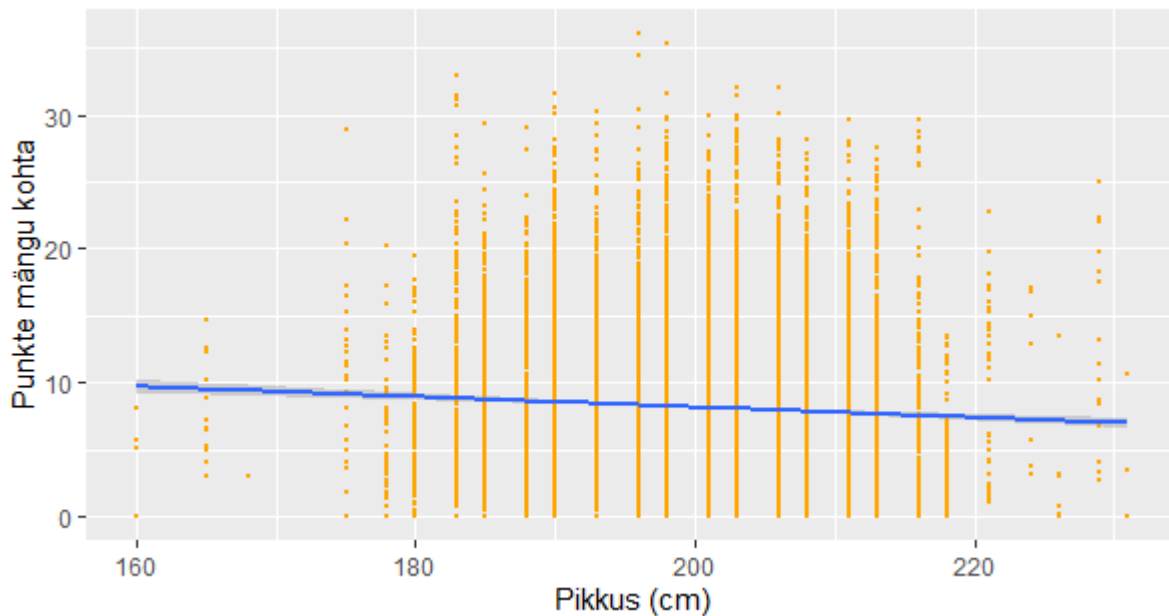
Nagu joonisel olevast regressioon sirgest näha, siis mida pikem on mängija, seda rohkem ta kaalub, mida oli ka arvata. Valdava enamus mängijate kaal jääb vahemikku 75kg kuni 125kg ja pikkus 180cm kuni 210cm. Kõige pikem mängija on olnud 231.14 cm ja kõige lühem 160.02.

Uurime nüüd ka, milline on kümne kõige populaarsema korvpalli ülikooli keskmine +- reiting (net_rating). Selle jaoks teeme karpdiagrammi ning filtreerime sealt välja väga ekstreemsed väärtused.



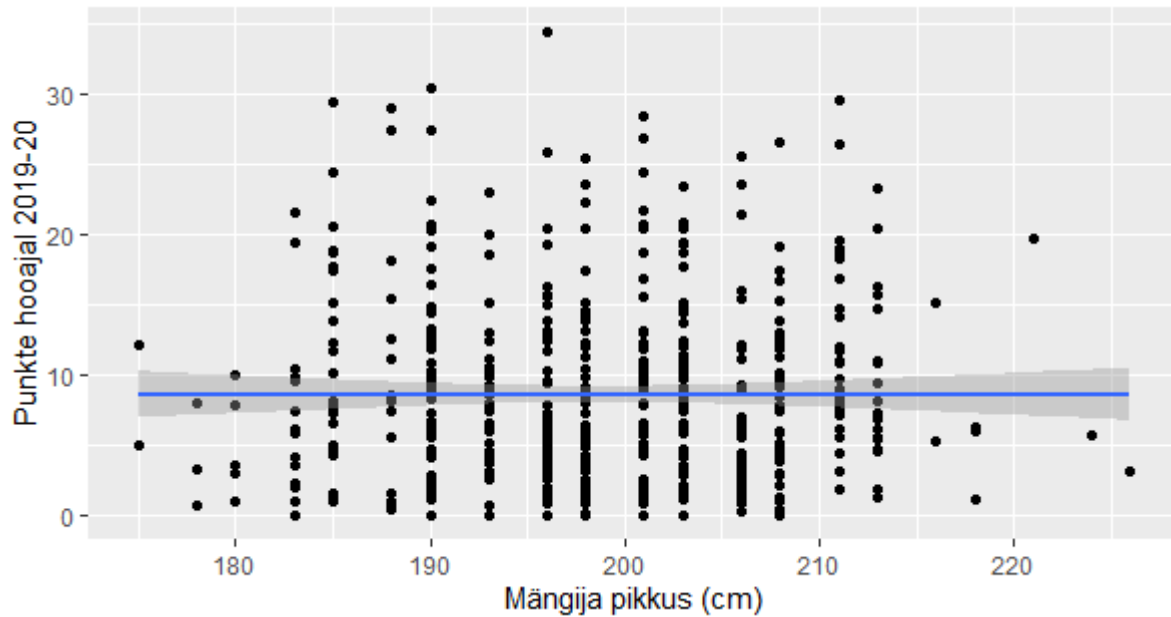
Joonis 4. Vastavate ülikoolide mängijate +- reiting 100 rünnaku ja kaitse kohta

Nagu näha, siis 10 kõige rohkem NBA mehi “tootnud” ülikooli vahel pole nende mängijate +- reitingus suuri erinevusi. Mediaan on kõigil üpris sarnane, ehkki pisut üllatav on see, et enamikul neist ülikoolidest jääb mediaan veidi alla nulli ehk 100 rünnaku/kaitse kohta on nende meeskond vastasmeeskonnale keskmiselt kaotanud. Samas võib seda mõjutada ka mõne mängija äärmiselt kehv +- reiting, mistõttu kogu reiting langeb.



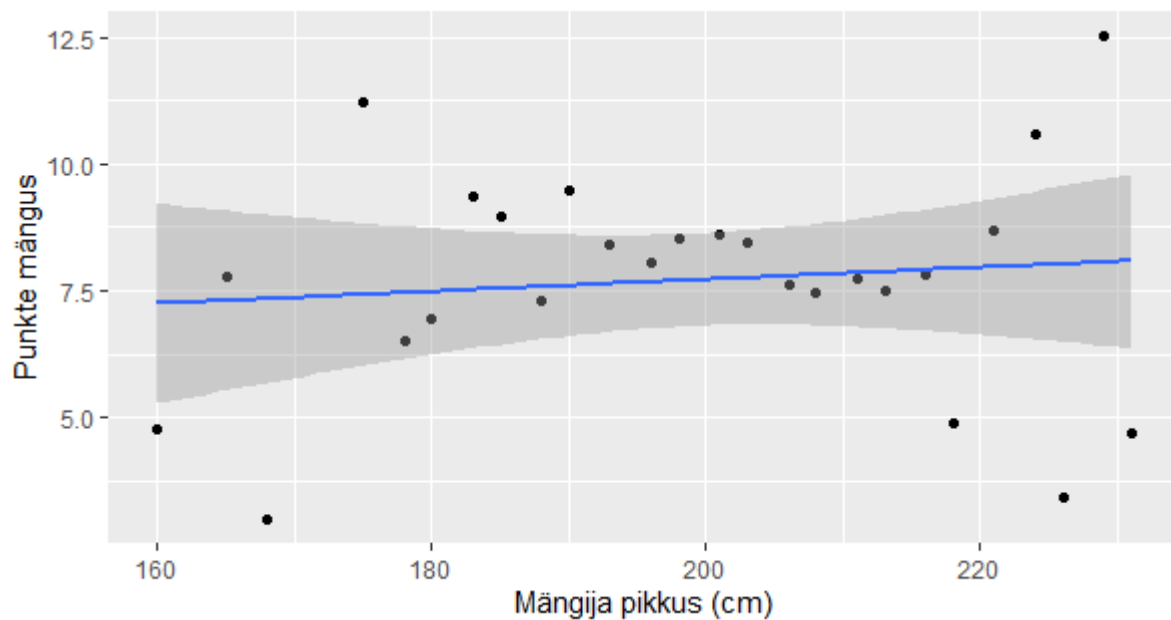
Joonis 5. Mängija pikkus ja punktide viskamine mängus 1996-2020

Joonisel 5 on näha, kui palju teatud pikkusega mängijad on punkte visanud. Lühemad mängijad viskavad pisut rohkem punkte kui pikad mängijad. Graafiku mõistes läheb lühikeste alla ka veel 183 näiteks, kus on juba palju kõrge skoorimisega hooaegu. Samas pikemas otsas (näiteks alates 215 cm) on mängijaid rohkem, kes ei visanud nii palju punkte ja see võib mõjutada regressioonisirget. Mõlemas otsas on näha äärmuslikke numbreid, ehk lühemad (alla 179 cm) võivad visata palju punkte ja väga pikad (üle 217 cm) võivad ka visata palju punkte. Kuid suurimad keskmised hooaja peale tulevad vahemikust 195-200 cm.



Joonis 6. Erineva pikkusega mängijate punktide viskamine mängus hooajal 2019-20

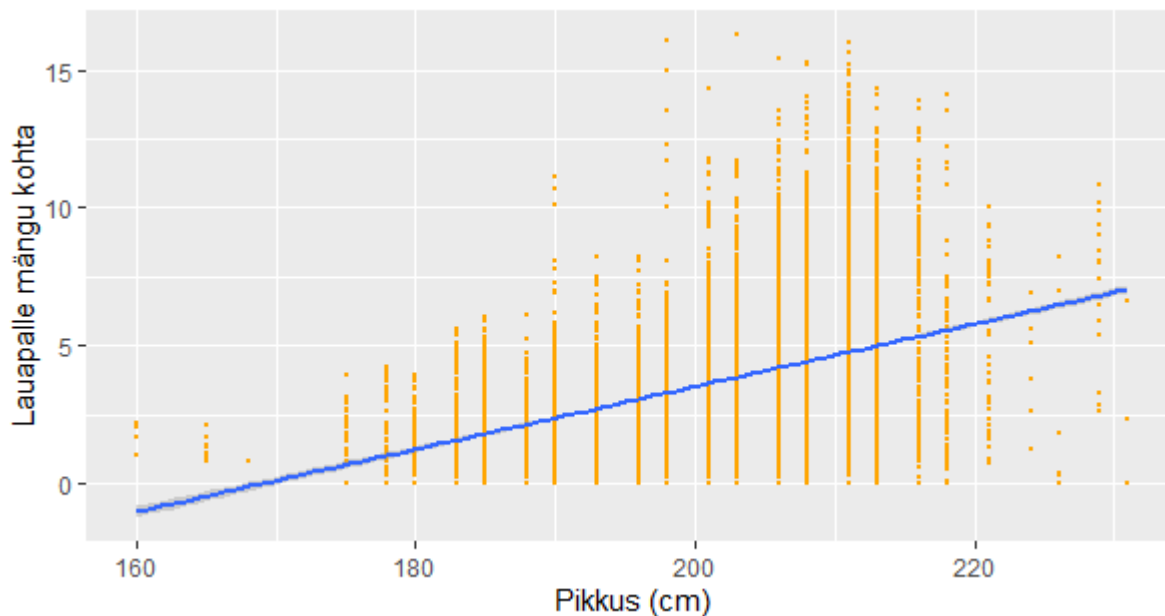
Hooegade lõikes ei ole väga suurt erinevust kujunenud, mis on pisut üllatav. Erinevad lihtsalt pisut selle tõttu, kas on mõni hea skoorija väga lühikeste või väga pikkade seas. Trendijoon on peaaegu alati pisut negatiivse tõusuga nagu joonisel 5, ehk lühemad viskavad veidi rohkem punkte. Erandiks on viimased 2 aastat, kus pikkusest endast ei olene punktide viskamine, näiteks joonis 6. See võib tulla sellest, et peamised staarid liigas on umbes 205 cm pikad, mis pikkuse mõttes on pikem kui keskmine. Lisaks on näha ainult 2 suuremat skoorijat, kes on alla 190 cm pikad. Praegu enam ei ole väga palju üle 213 cm pikkuseid mängijaid liigas, kes ei skoori nii palju.



Joonis 7. Mängija punktide viskamine pikkuste keskmiste järgi

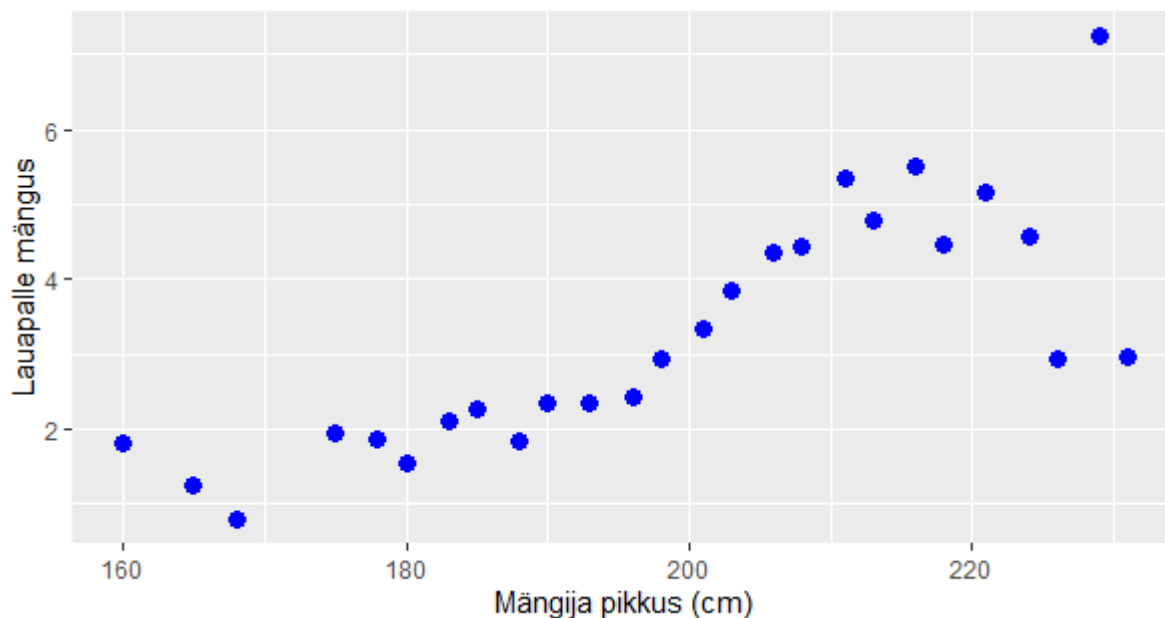
Kui vaadata pikkuste keskmisi, siis on varieeruvused suuremad. 180 cm kuni 210 cm on enam-vähem visanud keskmiselt 7 kuni 10 punkti mängus. Kui nendest pikkustest välja minna, siis numbrid hakkavad

varieeruma, kuna nende pikkuste puhul mõjutavad üksikud mängijad punktide hulka. Näiteks 175 cm puhul on 2 mängijat (Isaiah Thomas ja Nate Robinson) kogunud üle 10p mängus ja Thomas isegi korra 28 punkti mängus. Selle tõttu on juba standardhälve 175 cm mängijate puhul 6,71 cm.



Joonis 8. Erineva pikkusega mängijate lauapallide arv mängus

On näha, et pikemad mängijad võtavad keskmiselt rohkem laudu mängus. Siin on näha ühtlast kasvu ja et üle 10 lauapalli mängus hooaja peale on võtnud alla 2 meetrised mängijad üksikud korrad. Kõrgeim tulemus alla 190 cm pikkustel mängijatel on 6,1 lauda ja alla 180 cm mängijatel 4,2. Enim üle 12,5 lauapalli hooaja jooksul on võtnud 211 cm pikkused mängijad. Lauapallide keskmine on umbes 3,5 lauapalli mängus.



Joonis 9. Mängija lauapallide saamine mängus pikkuste keskmiste järgi

Kui vaadata pikkuste keskmisi, siis lauapallide keskmine arv tõuseb üpriski ühtlaselt pikkuse kasvades,

nagu oli oodata. Erandiks on graafiku otsad, aga seal on ka valim väike ja mõjutab üksiku mängija tase ja kui palju ta mänguaega näiteks sai.