

# Data-analyysin hyödyntäminen Counter-Strike: Global Offensive -pelissä

TURUN YLIOPISTO  
Tietotekniikan laitos  
TkK-tutkielma  
Tietotekniikka  
Kesäkuu 2025  
Matias Kokko

TURUN YLIOPISTO  
Tietotekniikan laitos

MATIAS KOKKO: Data-analyysin hyödyntäminen Counter-Strike: Global Offensive -pelissä

TkK-tutkielma, 20 s.  
Tietotekniikka  
Kesäkuu 2025

---

Tässä kandidaattitutkielmassa tarkastellaan, miten data-analytiikkaa voidaan hyödyntää Counter-Strike: Global Offensive (CS:GO) -pelissä. Tutkimuksessa selvitetään, miten dataa kerätään, jalostetaan tulkittavaan muotoon sekä hyödynnetään analyysityökaluissa ja -menetelmissä.

Tutkielmassa esitellään pelidatan keruun arkkitehtuuri, sekä jalostettua dataa hyödyntäviä analyysityökaluja. Näiden avulla on mahdollista tunnistaa pelitilanteita ja arvioida yksittäisen pelaajan tekemien päätösten vaikutusta pelin kulkuun. Nämä myös mahdollistavat pelin syvemmän ymmärryksen taktiseen näkökulmaan. Lisäksi on hieman pohdintaa miten CS:GO:n uudemman version Counter-Strike 2:n rakenteelliset muutokset ovat verrattavissa CS:GO analyysiin.

Tutkimuksen havainnoista havaittiin, että nykyiset tutkimukset keskittyvät enemmän pelien jälkeisiin analyysiin. Aiheesta siis todettiin aukko, siinä miten tätä dataa hyödynnetään taktiikoiden kehittämiseen ammattilaisotteluissa. Pelidatan tutkimisessakin on vielä kehitettävää, kuten koneoppimismallien hyödyntäminen. Näiden kehittämisestä olisi hyötyä erityisesti pelin ammattilaistasolla.

Asiasanat: Counter-Strike, data-analytiikka, e-urheilu

# Sisällys

<b>1</b>	<b>Johdanto</b>	<b>1</b>
1.1	Tutkielman tarkoitus, aiheen valinta ja tutkimuskysymykset . . . . .	2
1.2	Aineiston haku . . . . .	3
1.3	Tutkielman rakenne . . . . .	4
<b>2</b>	<b>Tausta</b>	<b>5</b>
2.1	E-urheilun ja Counter-Striken historia . . . . .	5
2.2	Counter-Strike -pelin rakenne . . . . .	7
2.3	Data E-urheilussa . . . . .	8
<b>3</b>	<b>Datan keräys ja jalostus</b>	<b>9</b>
3.1	Pelidatan tallentaminen ja sen formaatit . . . . .	9
3.2	Datan keruun automaatio ja arkkitehtuuri . . . . .	10
<b>4</b>	<b>Jalostetun datan hyödyntäminen</b>	<b>12</b>
4.1	ggViz - visuaalinen analyysityökalu . . . . .	12
4.2	Pelaajien ja pelirakenteen analyyseja . . . . .	13
<b>5</b>	<b>Pohdinta</b>	<b>15</b>
<b>6</b>	<b>Yhteenveto</b>	<b>19</b>
	<b>Lähdeluettelo</b>	<b>21</b>

# 1 Johdanto

Arjessa lähes kaikki jokapäiväiset askareemme voidaan kerätä dataksi. Dataa saadaan kerättyä monesta eri asiasta, esimerkiksi liiketoiminnasta, terveydestä ja turvallisuudesta. Tätä varten on kehittynyt datatiede, joka yhdistää data-analyysin, tilastotieteen sekä niiden muut menetelmät. Tämän avulla voidaan ymmärtää ja analysoida oikeita ilmiöitä kerätyn datan avulla. [1]

Data-analyysia käytetään urheilussakin, jotta joukkueet saisivat apua päätöksentekoihin. Päätöksentekoa ja mahdollisia hypoteeseja varten on kerättävä kaikki oleellinen, vanha ja tilastollinen tieto pelaajista. Tätä kutsutaan urheiluanalytiikaksi. Sen avulla on mahdollista arvioida joukkueen ja yksilön suorituksia, joukkueen mahdollisuuksia voittaa sekä monia muita asioita. [2] Esimerkiksi jalkapallossa on seurantalaitteita, joiden avulla voidaan kerätä tietoa pelaajien sijainneista, juoksumatkoista, syötöistä ja joukkueiden muodostelmista eri tilanteissa. Tästä kertyy todella paljon dataa analysoivataksi, myös turhaa dataa. Joten on tärkeää kehittää ratkaisuja, joiden avulla hyödyllinen data saadaan analysoitavaksi. [3]

Tämänkaltaista analyysia käytetään myös e-urheilussa (elektroniset urheilulajit) [4]. Perinteiset urheiluseurat hyödyntävät myös analytikoita parantaakseen strategisia päätöksiä, ja näitä samoja käytäntöjä E-urheilujoukkueet ovat alkaneet käyttämään hyödyksi. E-urheilun suosio on kasvanut viime vuosikymmenenä huimasti, vaikka edelleen käydään keskustelua siitä, että tulisiko se julistaa urheiluksi [5]. Kuitenkin on kehitetty jo lukuisia erilaisia sovelluksia, jotka keräävät E-urheilusta

saatavaa dataa. Näitä sovelluksia pelaajat, valmentajat sekä muut voivat käyttää hyödyksi kehittäessään erilaisia pelistrategioita [6]. Sovelluksia on myös tärkeää kehittää, sekä luoda uusia, sillä nykyisin e-urheilu on miljardiluokan markkina-ala. Vuonna 2020 digitaalisten videopelien markkina-arvo oli noin 126,6 miljardia USD. Tämä tutkimus kohdistuukin yhteen suosituimmista videopeleistä, Counter-Strike: Global Offensiveen, jota on myyty yli 25 miljoonaa kopiota. [4]

## 1.1 Tutkielman tarkoitus, aiheen valinta ja tutkimuskysymykset

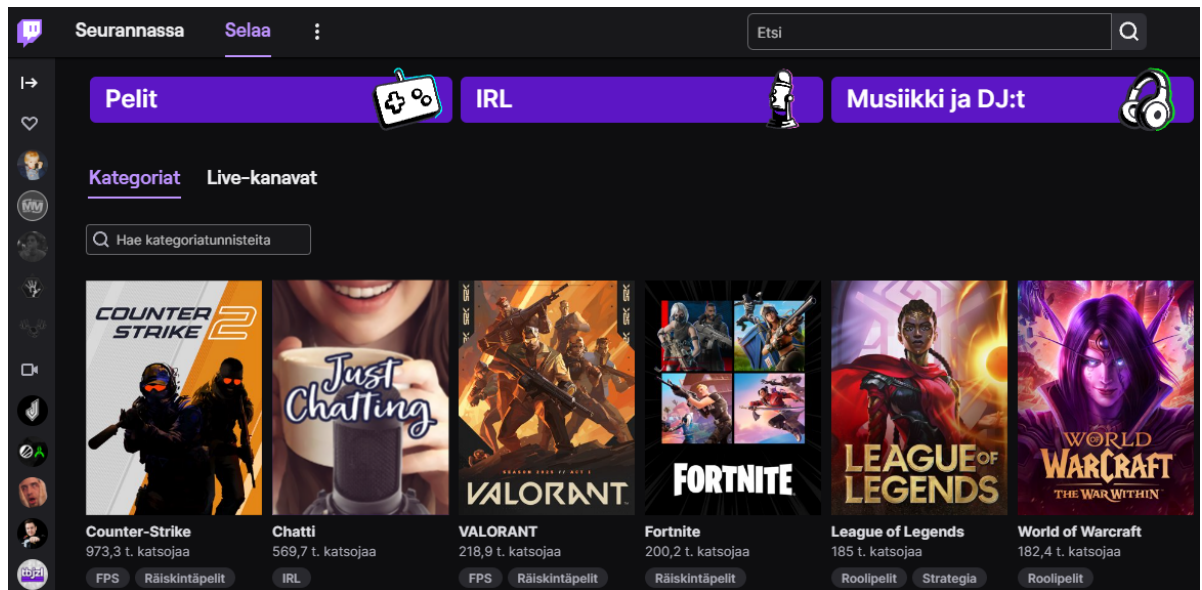
Tässä tutkielmassa tarkastellaan tapoja kerätä, jalostaa ja hyödyntää dataa Counter-Strike -pelisarjassa, tarkemmin rajattuna Counter-Strike: Global Offensive -pelissä (CS:GO). Vaikkakin CS:GO korvattiin vuoden 2023 loppupuolella Counter-Strike 2 -pelillä<sup>1</sup>, löytyy CS:GO:sta enemmän tietoa. Counter-Strike on myös erittäin suosittu pelisarja E-urheilussa, ja sen kuukausittaiset pelaajat yltyvät miljooniin<sup>2</sup>. Tämän työn tutkimuskysymyksiä ovat:

- Miten dataa kerätään ja miten se muutetaan hyödylliseksi?
- Mihin hyödyllistä dataa käytetään?

---

<sup>1</sup><https://www.esports.net/news/counter-strike/cs2-release-date/>

<sup>2</sup><https://www.counter-strike.net/cs2>



Kuva 1.1: Kuva otettu 9.2.2025 <https://www.twitch.tv/directory> -sivulta. Kuva kertoo, kuinka paljon katsojia eri peleillä on kyseisellä hetkellä. Kuvanotto aikana oli käynnissä IEM Katowice 2025 finaaliottelu Counter-Strike 2 -pelissä, jolloin kyseisen pelin katsojaluku saavutti melkein miljoonan.

## 1.2 Aineiston haku

Tähän tutkielmaan on haettu tietoa IEEE Xploresta sekä Google Scholarista. Hakua varten listattiin monia hakusanoja, kuten ("data analy\*", "data-analy\*", "Counter-Strike", "CS:GO", "CS2", "esports", "game analy\*", "match analy\*"). Näitä yhdistelemällä saatiin muodostettua hakulausekkeita, joiden avulla löydettiin tutkimuksia ja artikkeleita.

Kuitenkaan CS2-hakusanaan liittyen ei löytynyt vielä artikkeleita, joista olisi hyötyä tässä tutkielmassa. CS:GO-hakusanaan liittyen löytyi kahdeksan relevanttia tutkimusta, joita hyödynnettiin tässä tutkielmassa. CS:GO-hakusanalla löytyi myös vain muutamia tutkimuksia liittyen pelaajien psyykkiseen puoleen ja nettiyhtey-

den laadun vaikutusta peliin. Tutkielmassa on myös hyödynnetty sivustoja, kuten HLTV<sup>3</sup> ja Leetify<sup>4</sup>. HLTV on kilpailulliseen Counter-Strikeen keskittyvä sivusto [7], josta saatiin tilastodataa tutkimukseen, kuten HLTV 2.0 -arvioinnin selitys ja tiettyjen erien voittotodennäisyyksiä. Leetifyn avulla arkipelaaja voi analysoida omaa pelaamistaan [8], ja tutkielmassa hyödynnettiin Leetify:n tekemää blogia, jossa oltiin kerätty dataa pistoolieristä.

### 1.3 Tutkielman rakenne

Luvussa 2 aiheena on taustat. Taustoissa kerrotaan e-urheilun historiasta sekä Counter-Strike -pelisarjan historiasta. Luvussa esitellään myös miten Counter-Strike peli toimii sekä kerrotaan, miten dataa on hyödynnetty perinteisissä urheilulajeissa, kuten jalkapallossa ja miten sitä voidaan vastaavasti käyttää e-urheilussa.

Luvussa 3 kerrotaan pelidatan keräämisestä eli miten se tallennetaan, sekä kerätyn datan jalostuksesta tietynlaiseen formaattimuotoon. Luvussa 4 tarkastellaan jalostettua dataa. Siinä esitellään erilaisia analyysityökaluja ja -menetelmiä, sekä pohditaan hieman, että miksi ne edistäisivät Counter-Striken analysointia.

Luvussa 5 pohditaan havaintoja, joita nousi esille tutkielman aikana, kuten tutkimusaukkoja ja analyysien verrattavuutta CS2:n peliin. Ja luvussa 6 vedetään yhteen tutkielman oleelliset asiat.

---

<sup>3</sup><https://www.hltv.org/>

<sup>4</sup><https://leetify.com/>

## 2 Tausta

Tämä luku käsittelee e-urheilun taustoja. Luvun alussa tarkastellaan e-urheilun historiaa ja sen suosion kasvua. Sen jälkeen selitetään miten Counter-Strike pelisarja on saanut alkunsa ja minkälaisia eri vaiheita kyseisellä pelisarjalla on ollut. Kolmannessa alaluvussa kerrotaan Counter-Strike -pelin rakenteesta, jotta saataisiin parempi ymmärrys sen säännöistä ja mekaniikoista, mikä auttaa ymmärtämään data-analyysin hyödyntämistä kyseisessä pelissä. Ja viimeiseksi käsitellään lyhyt katsaus videopeleistä kerättävästä datasta.

### 2.1 E-urheilun ja Counter-Striken historia

Kilpailullisen e-urheilun katsotaan alkaneen joko *Nim*- tai *Tennis for Two* -pelistä. 1940 vuonna julkaistiin peli nimeltä *Nim*, jonka keskeisenä tekijänä oli kilpailuhenkiyys, sillä siinä kaksi pelaajaa kilpailivat toisiaan vastaan. *Nim* -peli oli myös luultavasti ensimmäinen turnauspeli, jossa oli katsojia mukana. Kuitenkaan kyseistä peliä ei tehty viihdytystä varten, ja tämän takia kilpailullisen e-urheilun toinen aloittaja oli *Tennis for Two* videopeli, joka julkaistiin vuonna 1958. Tässä kyseisessä pelissä otettiin huomioon myös viihdyttävyyttä, sillä peli sisälsi katsojan roolin ja sen avulla oli mahdollista seurata kahden pelaajan välistä kamppailua. [9]

E-urheilu ei silti ollut tuohon aikaan kannattavaa liiketoiminnallisesti, sillä tietokoneet maksoivat paljon. Tilanteeseen tuli kuitenkin muutos, kun tietokoneet halpenivat ja vuonna 1972 järjestettiin ensimmäinen E-urheiluturnaus, *Intergalactic*

*Spacewar Olympics*. Videopelien suosio kasvoi vielä voimakkaammin 1990-luvulla, kun henkilökohtaisista tietokoneista tuli vielä edullisempia. Toinen suosion kasvattaja oli internet, jonka myötä pelaajien oli mahdollista ylläpitää ja jakaa tuloksiaan muiden pelaajien kanssa eri maiden välillä [5].

Näiden kahden kehityksen avulla kilpailullisen pelaamisen suosio alkoi kasvaamaan. 1990-luvun lopulla ja 2000-luvun alussa nousi esiin muutamia organisaatioita, jotka järjestivät turnauksia, kuten Cyberathlete Professional League (CPL), Professional Gamers League (PGL), Deutsche Clanliga (DeCL), Electronic Sports League (ESL), French Electronic Sports World Cup (ESWC) ja Major League Gaming (MLG). Muutamat näistä organisaatioista ovat järjestäneet myös erittäin tunnettuja Counter-Strike -turnauksia<sup>1</sup>. 2000-luvulla e-urheilun suosio on kasvanut ja sen taloudellinen asema on vakiintunut. [9]

1990-luvun loppupuolella kaksi pelaajaa Minh Le ja Jess Cliffe tekivät modifikaation *Half-Life* -peliin, jonka omistaa Valve Corporation (Valve) ja tästä modifikaatiosta alkaa Counter-Striken historia.

Valve kiinnostui kyseisestä modifikaatiosta ja palkkasi modifikaation luoneet henkilöt kehittämään itsenäistä versiota tästä pelistä. Vuoden 2000 lopussa pelistä julkaistiin ensimmäinen oma versio, *Counter-Strike*. Vuonna 2003 pelistä julkaistiin uudempi versio *Counter-Strike 1.6*. Pian tämän jälkeen pelisarjasta julkaistiin kaksi uutta versiota, *Counter-Strike: Condition Zero* vuoden 2004 alussa sekä *Counter-Strike: Source* vuoden 2004 lopussa. Vaikka näissä kahdessa pelissä oli huomattavasti paremmat grafiikat, saivat nämä kaksi peliä silti paljon kritiikkiä Counter-Strike -yhteisöltä siitä, kuinka huonosti pelit olivat optimoitu. Tämä herätti erimielisyyttä ja jakoi Counter-Strike -yhteisön kahtia: toiset pysyivät *Counter-Strike 1.6* -pelissä ja toiset siirtyivät *Counter-Strike: Source* -peliin [9].

---

<sup>1</sup><https://www.hltv.org/>

Seuraavaa versiota pelisarjasta jouduttiin odottamaan useampia vuosia ja vuonna 2012 julkaistiin uusi versio, *Counter-Strike: Global Offensive* eli CS:GO. Sen yksi haasteista oli saada yhdistettyä aikaisemmin erilleen jakautuneet Counter-Strike -yhteisöt. Yksi suurimmista päivityksistä kyseiseen pelisarjaan tuli tämän version myötä, nimittäin pelejä oli mahdollista hakea oman taitotason mukaan.

Pelisarjan viimeisin versio, Counter-Strike 2, julkaistiin vuoden 2023 loppupuolella. Toisin kuin edellisissä versioissa, tämä päivitys tuli suoraan CS:GO:n päälle, eikä CS:GO ole enää pelattavissa.<sup>23</sup>

## 2.2 Counter-Strike -pelin rakenne

Counter-Strike: Global Offensive on ensimmäisen persoonan ammutapeli, jossa kamppailee keskenään kaksi viiden hengen joukkuetta. Pelissä itsessään on kaksi tiimiä, *Counter-Terrorists* (CT) sekä *Terrorists* (T). Kumpikin joukkue joutuu pelaamaan molemmilla puolilla, ja puolet vaihtuvat 15 erän jälkeen. Joukkue voittaa kartan, kun he ovat voittaneet 16 erää. Molemmat pelin puoliskoista alkavat pistoolikierroksella, jolloin kummallakin joukkueella on vain hieman rahaa ostaa tarvikkeita kuten suojaliiviä tai utiliteetteja, ei siis raskaampia aseita. Seuraavien kierroksien varallisuus korreloituu edellisten kierrosten menestyksen mukaan, voitetusta erästä saa enemmän rahaa kuin hävitystä erästä ja silloin voi ostaa parempia aseita.

Molemmilla joukkueilla on yksi sama keino saada erävoitto eliminoimalla kaikki vastapuolen pelaajat. Jokaisella pelaajalla on erän alussa saman verran elämäpisteitä, ja niiden loputtua kyseinen pelaaja eliminoiduu loppueräksi. Vahinkoa toiselle pelaajalle voi tehdä aseiden ja utiliteettien avulla, kuten käsikranaatilla.

Pelin kartoissa on kaksi pommialuetta, ja näitä alueita hyödyntäen molemmilla tiimeillä on eriävät keinot voittaa erä. T-tiimin tehtävänä on päästä jommallekum-

---

<sup>2</sup><https://www.youtube.com/watch?v=B35zWsp0r4Y>

<sup>3</sup><https://blog.acer.com/en/discussion/1371/the-history-of-counter-strike>

malle pommialueelle, asettaa pommi ja puolustaa sitä, kunnes se räjähtää. Täten CT-tiimin pitää puolustaa pommialuetta, ja jos T-tiimi saa pommin asetettua pommialueelle, voi CT-tiimi voittaa erän vielä purkamalla pommin. Jos erän aika loppuu, eikä pommia ole asetettu ja molemmissa tiimeissä on vielä pelaajia elossa, niin CT-tiimi saa erävoiton. [10][11]

CS:GO -turnauksissa pelikartat rajoittuvat vain seitsemään eri karttaan, ja otte-  
lut ratkaistaan usein paras kolmesta logiikalla. Eli molemmat joukkeet saavat valita  
ja poistaa karttoja kyseisistä seitsemästä kartasta, jotta jäljelle jää pelattavaksi kol-  
me karttaa. Tämä kasvattaa data-analyysin hyötyä, sillä jokaisen ammattilaisjouk-  
kueen voittoprosentit eri kartoista löytyvät HLTV-sivustolta.<sup>3</sup>

## 2.3 Data E-urheilussa

Datan hyödyntämistä on käytetty jo muissakin urheilulajeissa, kuten jalkapallossa.  
Kerätyn datan avulla on mahdollista luoda taktiikoita, joiden avulla voidaan saa-  
vuttaa tietty tavoite tai etu. Jalkapallossa taktiikoita on käytetty esimerkiksi siihen,  
miten joukkue hallitsee tilaa eli miten joukkue hallitsee kenttää joko hyökkäys- tai  
puolustustilanteessa. [3]

Tätä samaa tekniikkaa on onnistuttu käyttämään myös videopeleissä pelidata-  
analytiikan avulla. Sen avulla voidaan oppia, miten ihmiset ovat vuorovaikutuksessa  
pelin kanssa, ja sen myötä tunnistaa eri malleja, siitä miten eri tilanteisiin on rea-  
goitu. Kerätty data tallennetaan ja myöhemmin analysoidaan taktiikoita varten. [6]  
Datan kerääminen on tavallaan helpompaa e-urheilussa kuin perinteisissä urheilu-  
lajeissa, sillä tapahtumat tallennetaan suoraan palvelimelta, esimerkiksi lokitiedos-  
toihin.

## 3 Datan keräys ja jalostus

CS:GO -pelit toimivat asiakas-palvelin-arkkitehtuurilla, ja kilpailullisissa peleissä palvelin päivittyy 128 kertaa sekunnissa eli noin 8 millisekunnin välein. Jokaisen tikin aikana palvelimelle tallennetaan pelaajien syötteet eli näppäimistön painallukset ja hiiren klikkaukset, ja nämä sykronoidaan kaikille pelaajille myös jokaisen tikin aikana. Palvelin myös tallentaa, sekä päivittää kaikille pelaajille olennaiset tapahtumat - erän alun ja lopun tai pommin räjähdysten/purkamisen - näitä pelaajat eivät voi suoraan hallita. [10]

### 3.1 Pelidatan tallentaminen ja sen formaatit

Jotta CS:GO peleistä saataisiin kerättyä analysoitavaa dataa, täytyy hyödyntää kah- ta erilaista tietolähdetyyppiä: *korkean tason (engl. high-level)* ja *matalan tason (engl. low-level)* lähteitä.

*High-level data* on peleistä yleisesti kerättävää dataa. Tätä hyödyntävät verk- kosivustot, joista tunnetuin on HLTV. Ne keräävät dataa ammattilaispeleistä, - joukkueista sekä niiden pelaajista ja yksinkertaisista tilastoista. Tähän dataan si- sältyy muun muassa pelien lopputulokset ja pelaajien lopputilastot (*kills, assists, deaths*).

*Low-level data* on yksityiskohtaisempaa dataa. Se sisältää muun muassa tiedot pelaajien varusteista, sekä interaktioista, kuten liikkeen, sijainnin ja ampumisen ti-

kin tarkkuudella. Tämä data tallennetaan jokaisesta pelistä niin kutsuttuun *demo-tiedostoon*.

*Demo-tiedostojen* tulkitsemiseksi, data pitää muuttaa rakenteettomasta lokitiedostosta parserin *esim. JSON-formaatin* avulla rakenteelliseen muotoon. Jotta rakenteellisesta muodosta saataisiin helposti ymmärrettävä ja luettava, täytyy parserin jakaa pelit pienempiin tilanteisiin, jotka sisältävät matalan tason lähteiden yksityiskohtaiset tiedot. [10]

## 3.2 Datan keruun automaatio ja arkkitehtuuri

Tutkimuksessa *A Practical Proposal for Building Counter-Strike Datasets* esittelee arkkitehtuuri, jonka avulla on mahdollista kerätä CS:GO-dataa. Sen avulla voi myös jäsentää, sekä tallentaa kyseistä dataa. Arkkitehtuuri rakentuu viidestä komponentista, jotka ovat *hakurobotti*, *demo tiedostojen lataaja*, *parseri*, *orkestroija* ja *indeksoija*. Näiden tehtäviä arkkitehtuurissa ovat:

**Hakurobotti** Hakee HLTV -sivustolta pelin, josta se kerää *high-level datan*, sekä linkin *demo-tiedoston* lataamiseen. Näistä se luo manifest-tiedoston.

**Demo-tiedoston lataaja** Lataa *demo-tiedoston* manifest-tiedoston linkkien avulla ja tallentaa *demo-tiedoston*.

**Parseri** Purkaa *demo-tiedoston* CSV-tiedostoksi.

**Orkestroija** Muuntaa CSV-tiedoston Parquet-muotoon.

**Indeksoija** Tekee SQLite-pohjaisen hakutietokannan, johon se indeksoi *low-level dataa* analysoinnin nopeuttamista varten.

Tiedostot muutetaan lopuksi Parquet-muotoon, sillä sen avulla on mahdollista lukea vain tarvittavat osat. Lisäksi se vie paljon vähemmän levytilaa kuin CSV-

tiedosto. Vertailun vuoksi, kun CSV-tiedosto vie 83,2 GB levytilaa, niin Parquet-muotoiltu vie vain 9,86 GB levytilaa. [4]

Esimerkiksi tämänkaltaisesti muodostetun jalostetun datan avulla voidaan kehittää analyysityökaluja ja -menetelmiä, joiden avulla raakadataa voidaan tutkia ja hyödyntää eri tarkoituksiin, kuten taktiikoiden purkamiseen, pelitilanteiden toistamiseen ja pelaajien toiminnan arviointiin. Tämänkaltaisten työkalujen avulla on mahdollista ymmärtää syvällisemmin pelin kulkua ja se auttaa valmentajia, analytikoita sekä arkipelaajien päätöksenteossa.

Tutkimuksessa *ggViz: Accelerating Large-Scale Esports Game Analysis* aloitettiin kehittämään analyysijärjestelmää, ja sitä varten haastateltiin neljää ammattiesports-toimijaa (valmentajia, analytikoita ja managereita). Heistä kolme kuului tuolloin CS:GO joukkueorganisaatioihin, jotka olivat maailman listalla kymmenen parhaan joukossa ja yksi heistä oli joukkueorganisaatioissa listalla 50 parhaan joukossa. Moni heistä korosti, että analyysityö on hidasta ja sopivia pelitilanteita on vaikea löytää. Joten analyysijärjestelmän suurimmaksi prioriteetiksi tuli eri pelitilanteiden löytäminen nopeasti — erityisesti eri varustetasoilla pommialueisiin hyökkääminen, sekä niiden puolustaminen. [10]

## 4 Jalostetun datan hyödyntäminen

Tämän luvun tarkoituksena on esitellä kehittyneempiä järjestelmiä. Yksittäisistä otteluista kertyy paljon dataa, ja niiden manuaalinen läpikäyminen ja analysointi vie paljon aikaa. E-urheilussa pelianalyysien tarpeet ovat suuret, ja ammattilaistasolla joukkueet tarvitsevat tehokkaampia järjestelmiä keräämään ja analysoimaan dataa.

### 4.1 ggViz - visuaalinen analyysityökalu

CS:GO-otteluiden jälkianalyysi vie paljon aikaa manuaalisesti, sillä pelit kestävät kauan ja ne on käytävä alusta loppuun, jotta löydettäisiin sopivia tilanteita. On kuitenkin olemassa työkaluja, joiden avulla on mahdollista katsoa pelitallenteita, mutta sellaisia työkaluja ei juurikaan ole, joiden avulla olisi mahdollista hakea tietynlaisia pelitilanteita. Tämä nopeuttaisi ja parantaisi analyysiprosesseja. Tämänkaltaista teknologiaa on onnistuttu hyödyntämään perinteisissä urheilulajeissa. E-urheilu ei ole saavuttanut vielä samanlaista analytiikka kehitystasoa kuin perinteiset urheilulajit. Yksi haaste on se, että datalle ei ole kehitetty yhteistä vakiintunutta rakennetta, joka vaikeuttaa erilaisten analyysijärjestelmien kehittämistä. [10]

Tutkimuksessa *ggViz: Accelerating Large-Scale Esports Game Analysis* esitellään ratkaisuksi ggViz-analyysijärjestelmä. Sen avulla pystyttäisiin hakemaan tiettyjä pelitilanteita automaattisesti suuresta CS:GO-otteluaineistosta. Analyysijärjestelmän avulla on mahdollista muunta pelaajien sijainnit token-muotoisiksi kuvauksiksi, sekä hakukysely voi tapahtua koko kartan kattavista sijainneista tai vain osittaisista

aluekyselyistä. Tämänkaltainen järjestelmä nopeuttaisi juuri oikeiden pelitilanteiden löytämistä, sekä tehostaisi taktiikoiden analysointia. Järjestelmää voi myös laajentaa muihinkin e-urheilupeleihin. [10]

## 4.2 Pelaajien ja pelirakenteen analyyseja

Toisessa tutkimuksessa *Valuing Player Actions in Counter-Strike: Global Offensive* esiteltiin kolme asiaa, joiden avulla on myös mahdollista analysoida pelidataa. Ensimmäinen avain asia oli, että peleistä tehtäisiin yhteinäisiä tietomalleja. Tämän avulla pelaajien tekemät interaktiot voitaisiin tallentaa systemaattisesti analysoitaviksi. Tutkimuksessa tutkijat kehittivät avoimen lähdekoodin työkalun, jonka muutenkin tutkijat ja analyysityökalut voivat hyödyntää. [11]

Lisäksi he kehittivät mittarin pelaajien etäisyyksien analysointiin. Sen avulla voidaan mitata yksittäisten pelaajien välisiä etäisyyksiä. Kyseisen mittarin avulla voidaan löytää tietynlaisia avainpiirteitä — kuten ratkaisevilla hetkille pelaajien etäisyydet toisistaan ja pelaajien sijainnit kartalla eri pelitilanteissa. Näiden avulla on mahdollista tutkia miten yksittäisen pelaajan asema voi vaikuttaa peliin. [11]

Viimeisenä kehityksenä he esittivät niin kutsutun WPA-mallin. Tällä mallilla on mahdollista mitata yksittäisen pelaajan vaikutuksia erän lopputulokseen. Sen avulla voidaan siis arvioida eri interaktioiden vaikutusta joukkueen voittotodennäköisyyteen. Kyseisiä interaktioita on mm. eliminaatio tai valokranaatin heitto. Malli ottaa huomioon eri muuttujia pelitilanteissa, kuten elossa olevien pelaajien määrän ja erän jäljellä olevan ajan, jonka jälkeen se vertailee pelitilannetta ennen ja jälkeen interaktion. Yleensä pelaajat arvioidaan vain kootuista tilastoista pelien jälkeen, mutta tämän mallin avulla on myös mahdollista arvioida yksittäisten tilanteiden hyötyä. [11]

HLTV Rating 2.0 on erittäin tunnettu analyysimenetelmä etenkin ammattilaispelien keskuudessa. Sen avulla voidaan laskea kullekin pelaajalle arvosana pelin

lopuksi. Se ottaa huomioon pelaajan tapot, kuolemat, tehdyn vahingon sekä eräkohtaiset teot, esim. jos pelaaja onnistuu sokaisemaan vastustajan valokranaatilla ja oman tiimin pelaaja onnistuu eliminoimaan vastustajan tämän ollessa sokea, kasvaa valokranaatin heittäneen pelaajan HLTV Rating 2.0.<sup>1</sup> [11]

Kuitenkaan tilastot eivät aina suoraan kerro pelin kulusta, vaan siihen on voinut vaikuttaa satunnaisia muuttujia. Tutkimuksessa *Analysis of Counter-Strike: Global Offensive* tutkittiin miten CS:GO -peleissä taito ja sattuma ovat tasapainossa. Data kerättiin virallisista turnausotteluista kuin myös tavallisista verkkopeleistä. Aineistoista saadut tulokset antoivat kaksi lopputulemaa. Pelaajien pelimekaniikkataitotaso oli vahvassa roolissa pelin kulkuun, mutta myös satunnaisuudet, kuten yksittäisten kierrosten tapahtumat saattoivat muuttaa pelien lopputulosta. [12] Tämä tekee peleistä mielenkiintoista seurattavaa, sillä pelit eivät ole täysin ennustettavissa.

Lisäksi tutkimuksessa mainitaan, että pelin sisäisiä kartoja ja pelin ominaisuuksia on muutettu ajan saatossa, jotta pelien pelirakennedata pysyisi tasapainossa. [12] Kyseisiä ominaisuuksia on esim. eri aseiden hinnat tai Counter-Terroristien ja Terroristien aloituspisteet kartoissa.

---

<sup>1</sup><https://www.hltv.org/news/20695/introducing-rating-20>






## 5 Pohdinta

Tutkimuksen aikana havaittiin, että data-analyysin hyödyntämistä CS:GO -pelissä oltiin tutkittu vain datan jalostuksen ja siitä tehtyjen analyysimenetelmien kannalta. Tätä aihetta ei ollut tutkittu siltä kannalta, että miten jalostettua dataa voitaisiin hyödyntää taktisesta näkökulmasta esimerkiksi ennen ammattilaispelejä. Monet tutkimukset keskittyivätkin siihen, miten pelaajien tekoja voidaan arvioida pelien jälkeen, mutta ei siihen, miten datasta voitaisiin saada hyödyllistä etua ennen pelejä.

Taktisesta suunnittelusta tehtyjen tutkimusten puutos oli huomattava aukko tässä aiheessa. Erityisesti ammattilaispelejä varten valmistaudutaan etukäteen analysoimalla vastustajan pelaamista ja kehittämällä sitä vastaan omia taktiikoita. Pelianalyysijä tehdään myös edelleen manuaalisesti valmennustiimien ja analyttikoiden toimesta, mikä on aikaa vievää. Jatkotutkimuksena voisikin olla tekoälyn hyödyntäminen datan keruussa. Koneoppimismalleja voisi kouluttaa tunnistamaan samankaltaisia tilanteita sekä strategisia virheitä, jotka eivät ole olleet optimaalisia yksittäisen kierroksen voittomahdollisuuksiin.

Counter-Strike -pelisarja on myöskin saanut uuden version CS:GO:n tilalle, CS2. Pelien mekaniikat ovat lähes samanlaiset lukuunottamatta uudistettuja grafiikoita sekä savukranaattien uutta ominaisuutta. Kuitenkin suurin ero, joka on vahvasti sidonnainen taktisiin suunnitteluihin, on erien määrät. CS:GO:ssa pelattiin niin kutsutussa MR15-muodossa (15 erää molemmin puolin), kun taas CS2:seen muutettiin MR-12-muotoon (12 erää molemmin puolin). Tässä uudessa formaatissa korostuu

virheiden painoarvo, sillä virheille jää vähemmän tilaa erien vähentymisen vuoksi. Etenkin pistoolierien voitto korostuu tässä uudessa formaatissa, koska jos sen voittaa niin suurella todennäköisyydellä voittaa myös seuraavankin erän <sup>1</sup>.

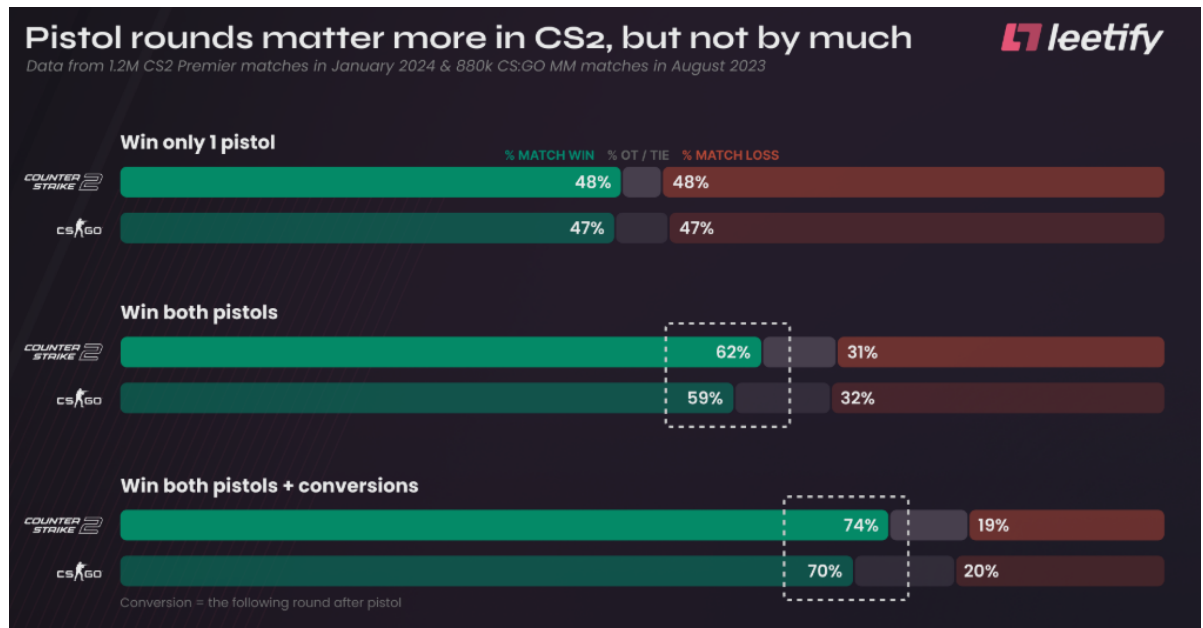
Team	Maps	Won - Lost	Pistol win %	Round 2 conv	Round 2 break
 G2	42	53 - 31	63.1%	81.1%	19.4%
 3DMAX	45	51 - 39	56.7%	76.5%	13.2%
 Spirit	31	37 - 25	59.7%	94.6%	8.0%
 MOUZ	74	79 - 69	53.4%	82.1%	17.6%
 Liquid	35	36 - 34	51.4%	75.0%	20.6%

Kuva 5.1: Kuva otettu 13.5. HLTV -sivustolta <https://www.hltv.org/stats/teams/pistols?startDate=2025-02-13&endDate=2025-05-13&rankingFilter=Top20>. Kuvassa on lajiteltuna viisi parasta joukkuetta pistoolierävoittojen mukaan viimeisen kolmen kuukauden ajalta. *Round 2 conv* on voittotodennäköisyys seuraavalle kierrokselle, jos he ovat voittaneet edeltävän pistoolierän ja vastaavasti *Round 2 break* on voittotodennäköisyys seuraavalle kierrokselle, jos he ovat hävinneet edeltävän pistoolierän.

Luvussa 1.2 mainittiin Leetify, joka on siis työkalu, jonka avulla arkipelaajatkin voivat tutkia omaa suoriutumistaan peleissä. Leetify on koonnut tallentamistaan peleistä dataa, siitä kuinka suurelle todennäköisyydellä voittaa pelin, jos voittaa pistoolikierroksia.

---

<sup>1</sup><https://www.hltv.org/stats/teams/pistols?startDate=2025-02-13&endDate=2025-05-13&rankingFilter=Top20>



Kuva 5.2: Kuva otettu 13.5. Leetify:n -blogisivustolta <https://leetify.com/blog/cs2-vs-csgo-pistol-rounds/>.

Leetify siis vertaili pistoolierien merkistystä molemmissa peleissä, CS:GO:ssa ja CS2:ssa. Kuten kuvasta 5.2 huomataan, niin pistoolierien voittojen merkitys on kasvanut muutamalla prosenttiyksiköllä CS2:ssa. Kyse ei ole siis vain yhden erän voitosta, kun voittaa pistoolierän, sillä sen voitettua joukkue saa huomattavan rahallisen edun seuraavalle erälle. Kyseinen tilasto ei ole kerätty pelkästään ammattilaispeleistä, vaan myös arkipelaajien peleistä. Arkipelaajien peleissä tämä saattaa korostua enemmän, sillä heillä ei ole tukena taktista valmennusta, jolla olisi pohdittuna palautumista rahallisesta tappiosta. Joten automaattisesti samankaltaisia tilanteita keräävästä analyysityökalusta olisi hyötyä arkipelaajallekin, sillä manuaalinen taktiikoiden kehittäminen vie erittäin paljon aikaa.

CS:GO:n -ammattilaispeleistä puuttuu myös yksi kiinnostava näkökulma, nimittäin live-ennustemallit. Tutkimuksessa *Win Prediction in Multiplayer Esports: Live Professional Match Prediction* tutkittiin live-ennustamisen hyödyntämistä e-

urheilulähetyksissä. Sen avulla voitaisiin informoida katsojiakin paremmin tietyissä pelitilanteissa. Kyseinen ennustus voi kuitenkin antaa väärää tuloksia ammattilaispeleissä, sillä aiemmat mallit ovat opetettu laajalla aineistolla kaiken tasoista peleistä. [13]

Tutkimuksen päätavoitteena oli siis saada valmistettua malli, jonka avulla voitaisiin ennustaa pelien voittotodennäköisyyksiä reaaliaikaisesti MOBA-pelissä. Kaikissa peleissä ei ole välttämättä selkeitä pisteitä, ja se voi satunnaisella katsojalle olla haastavaa tulkita pelin etenemistä. Tutkimuksessa tarkasteltiin myös kuinka suuri vaikutus ammattilaispelien ennustamiseen on, kun mallia koulutetaan e-ammattilaispelien datalla verrattuna pelkästään ammattilaispelien datalla. Kyseisen tutkimuksen tehtyä mallia käytettiin myös kansainvälisessä turnauksessa, sekä se oli ensimmäinen akateeminen työ, jota käytettiin e-urheilutapahtuman live-lähetyksessä. [13]

Vaikkakin MOBA-pelit eroavat CS:GO kaltaisista FPS-peleistä, voitaisiin kyseistä logiikka hyödyntää kehittämään myös FPS-peleihin reaaliaikaista voittotodennäköisyysmallia. Se voisi kasvattaa katsojien määrää entisestään, sillä vähemmän peliä seuranneet ymmärtäisivät eri asioiden vaikutuksen pelin kulkuun. Sekä se voisi olla lisäelementti live-lähetyksien selostajien selotukseen.

## 6 Yhteenveto

Tässä tutkielmassa tarkasteltiin, miten data-analyysiä voidaan hyödyntää Counter-Strike: Global Offensive -pelissä, sekä sitä miten peleistä kerätään dataa ja miten se voidaan jalostaa käyttökelpoiseen muotoon. Tutkielmassa käsiteltiin analyysimenetelmiä sekä -työkaluja, joiden avulla on mahdollista arvioida pelaajien suorituksia datan avulla.

Peleistä kertyvä data voidaan jakaa kahteen luokkaan, korkeaan tasoon sekä matalaan tasoon. Korkean tason data on tilastot, joita saa ammattilaispeleistä HLTV -sivustolta. Matalan tason data on yksityiskohtaisempaa, sillä se sisältää jokaisen pelaajan tekemän interaktion pelissä. Datan kerääminen on viety teknisesti pitkälle, ja tutkimuksessa esitellään tutkimuksen *A Practical Proposal for Building Counter-Strike Global Offensive Datasets* arkkitehtuuri, jonka avulla voidaan automaattisesti kerätä ja jäsenellä dataa.

Analyysityökalut sekä -menetelmät hyödyntävät tätä jalostettua dataa. Tämän avulla voidaan arvioida yksittäisiä toimintoja ja miten ne vaikuttavat peliin kuten *WPA-malli*, sekä myöskin voidaan hakea samankaltaisia pelitilanteita (*esim. ggViz*). Menetelmien avulla voidaan analysoida pelitilanteita ja löytää avainpareja, jotka toistuivat tai erottuivat joukosta. Vaikkakin työkaluja ja menetelmiä on kehitetty, ne keskittyvät pelien jälkianalyysiin. Tutkimuksen aikana aukko siinä, miten taktista suunnittelua tehdään ammattilaistasolla ennen pelejä.

Tutkimuksen pohdinta osiossa mietittiin myös CS:GO:n päivitetyn version CS2:n uuden formaatin vaikutusta peliin. Erien vähenemisen vuoksi pistoolierien painoarvo on noussut, joten niihin valmistautuminen korostaa analyysin ja valmistautumisen merkitystä. Lisäksi huomattiin, että MOBA-peleissä on hyödynnetty liveennustemalleja, jotka kertovat vastakkain pelaavien joukkueiden voittomahdollisuuksista. Tätä ei kuitenkaan ole ollut vielä CS:GO:ssa. Voidaan siis todeta, että data-analytiikan rooli laajentuu e-urheilussa, sillä eri pelien data-analytiikkaa voitaisiin hyödyntää muissakin peleissä. On myös mahdollisuuksia kehittää erilaisia koneoppimismalleja, jotka nopeuttaisivat datan keräämistä ja suodattamista. Tämä nopeuttaisi analyysiprosesseja myös ammattilaisjoukkueissa, sekä olisi hyvä kehittää sovelluksia myös arkipelaajille - kuten Leetify on jo osittain tehnyt.

# Lähdeluettelo

- [1] I. H. Sarker, "Data science and analytics: an overview from data-driven smart computing, decision-making and applications perspective", *SN Computer Science*, vol. 2, nro 5, s. 377, 2021.
- [2] P. HV, C. Nandini et al., "A Survey on the application of Data Science And Analytics in the field of Organised Sports", *arXiv preprint arXiv:2209.07528*, 2022.
- [3] R. Rein ja D. Memmert, "Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science", *SpringerPlus*, vol. 5, s. 1–13, 2016.
- [4] E. Rocha, H. Maio, D. S. Menasché ja C. Miceli, "A Practical Proposal for Building Counter-Strike Global Offensive Datasets", 2021.
- [5] J. Bousquet ja M. Ertz, "eSports: historical review, current state, and future challenges", teoksessa *Handbook of Research on Pathways and Opportunities into the Business of Esports*, IGI Global, 2021, s. 1–24.
- [6] R. P. F. Roque, "Esports–Video Game Data Analysis", tohtorinväitöskirja, 2023.
- [7] HLTV. "HLTV.org". Viitattu 22.6.2025. (2024), url: <https://www.hltv.org/>.
- [8] Leetify. "Leetify". Viitattu 22.6.2025. (2024), url: <https://leetify.com/>.

- 
- [9] T. M. Scholz ja T. M. Scholz, ”A short history of esports and management”, *eSports is Business: Management in the World of Competitive Gaming*, s. 17–41, 2019.
- [10] P. Xenopoulos, J. Rulff ja C. Silva, ”GgViz: Accelerating large-scale esports game analysis”, *Proceedings of the ACM on Human-Computer Interaction*, vol. 6, nro CHI PLAY, s. 1–22, 2022.
- [11] P. Xenopoulos, H. Doraiswamy ja C. Silva, ”Valuing player actions in counter-strike: Global offensive”, teoksessa *2020 IEEE international conference on big data (big data)*, IEEE, 2020, s. 1283–1292.
- [12] M. N. Rizani ja H. Iida, ”Analysis of counter-strike: Global offensive”, teoksessa *2018 international conference on electrical engineering and computer science (ICECOS)*, IEEE, 2018, s. 373–378.
- [13] V. J. Hodge, S. Devlin, N. Sephton, F. Block, P. I. Cowling ja A. Drachen, ”Win prediction in multiplayer esports: Live professional match prediction”, *IEEE Transactions on Games*, vol. 13, nro 4, s. 368–379, 2019.