

All onödig fakta kring Game of Thrones!

Filip Axelsson (@filipaxelsson)

2020-01-15

Introduktion

De flesta av oss har följt Game of Thrones slaviskt oavsett om det är böckerna eller om det är tv-serien. Så fort man har blivit fäst vid en huvudkaraktär så dödar dom personen och detta uppreppas flera gånger. I detta projekt kommer jag ge dig som läsare all onödig iformation du behöver om Game of Thrones, det kan vara allt ifrån vad är livslängden för respektive hus, hur många karaktärer finns det endast i böckerna och endast i serien, hur ser fördelningen ut i kön, vilket hus är störst, vilket hus har det dött flest personer i, och mycket mer onödig information. Men som man brukar säga "kunskap är inte tungt bära med sig" och vem vet kanske kommer du få en miljonkronorsfråga om just detta.

Data

Det finns säkert flera olika sätt att få tillgång till information kring Game of Thrones, jag har dock valt mig att använda en API, "<https://anapioficeandfire.com>". Jag kommer utgå ifrån att denna API har korrekt information till de säsonger och böcker som finns listade. API:n är uppbyggd på följande sätt, du hämtar hem data för böckerna genom en url-länk till exempel för första boken, "<https://anapioficeandfire.com/api/books/1>", detta genererar en massa olika variabler, bland annat en variabel "character" som i sin tur består av flera url-länkar, se "Tabell 1". Genom att använda respektive url-länk i "characters" kan ytterliggare information hämtas hem, men om ytterliggare information behövs som till exempel vilket hus karaktären tillhörde så måste ännu en url-länk till API:n användas. Denna princip följer genom hela API:n.

Table 1: Tabell 1: Exempel på struktur för variabeln "characters"

characters
https://anapioficeandfire.com/api/characters/2
https://anapioficeandfire.com/api/characters/12
https://anapioficeandfire.com/api/characters/13
https://anapioficeandfire.com/api/characters/16
https://anapioficeandfire.com/api/characters/20

Genom att skapa en funktion som läser in alla url-länkar går det smidigare att läsa in all data. Ett annat problem som uppstod är att vissa variabler innehåller text, så för att få ut den sökta informationen skapades även här en funktion som tog ut den sökta informationen och transformerade den. Ett exempel på en sådan variabel är "born", det vill säga när karaktären är född. Denna variabel består med en massa text följd med ett årtal där "AC" eller "BC" står efter. Jag skapade en funktion som tog ut den sökta informationen i "born", det vill säga året och om det stod "AC" så tas endast året ut, står det dock "BC" läggs ett minustecken till före, detta görs även för variabeln "dead". Jag gör detta eftersom vi då kan ta absolutbeloppet av "dead - born" för att få ut åldern på karaktären. Vi ska även notera att i denna datamängd/API finns det 10 stycken olika böcker samt 6 säsonger av tv-serien.

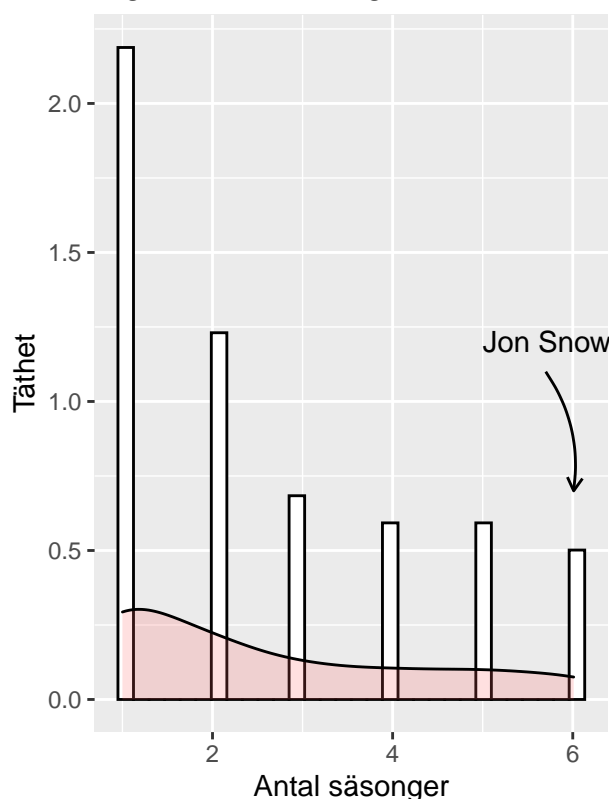
Analys av data

Förväntad “tid” i rampljuset för en karaktär

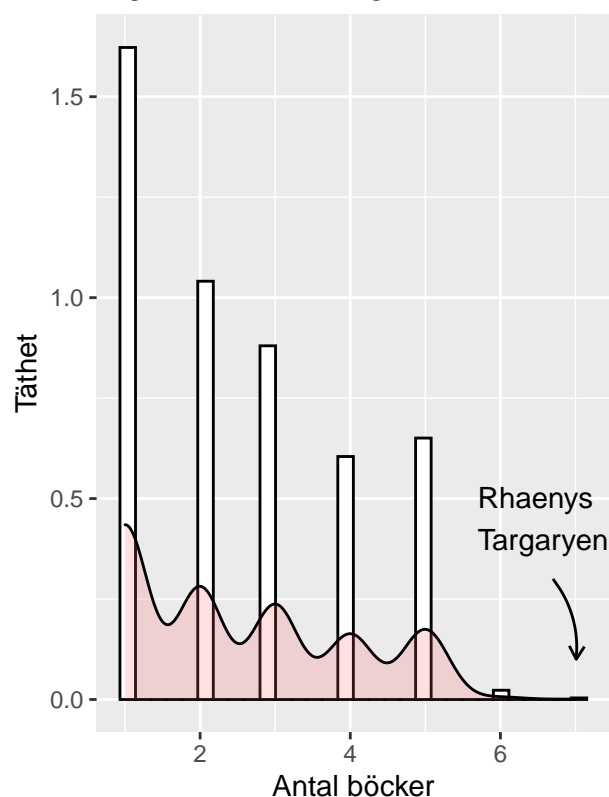
Genom att analysera “Figur 1” och “Figur 2” som visar hur många böcker respektive säsonger av tv-serien en karaktär är med i. Det går att dra en slutsats att antalet böcker och säsonger verkar vara exponentialfördelad. Vi kan utifrån detta se att det är väldigt få karaktärer som är med i alla säsonger i tv-serien, till exempel är “Jon Snow” en utav karaktärerna som är med i alla 6 säsonger. Om vi istället riktar fokus till böckerna så ser vi att ingen utav karaktärerna finns med i alla 10 böcker, den är även väldigt få som finns med i bara 6 stycken, det är färre än vad det var för tv-serien, dock den som har varit med i flest böcker är “Rhaenys Targaryen” som har varit med i 7 böcker.

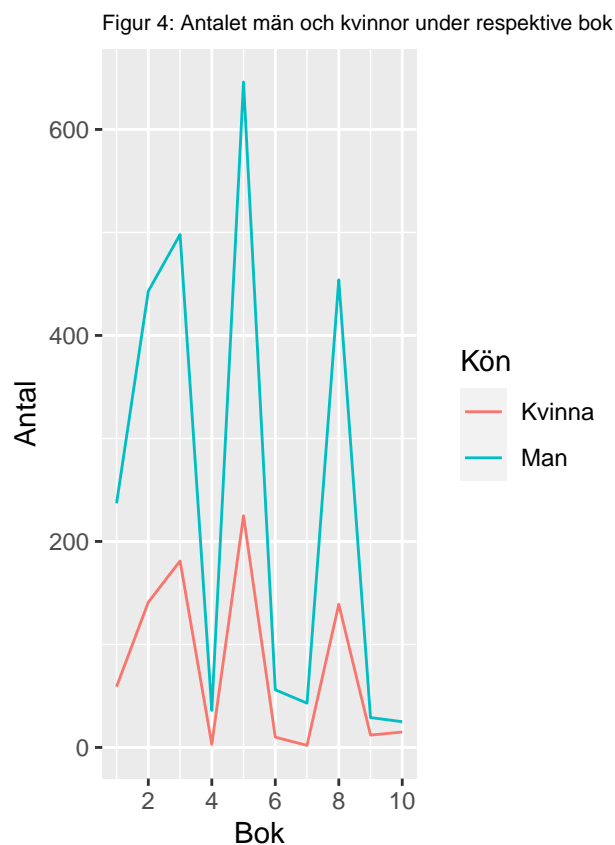
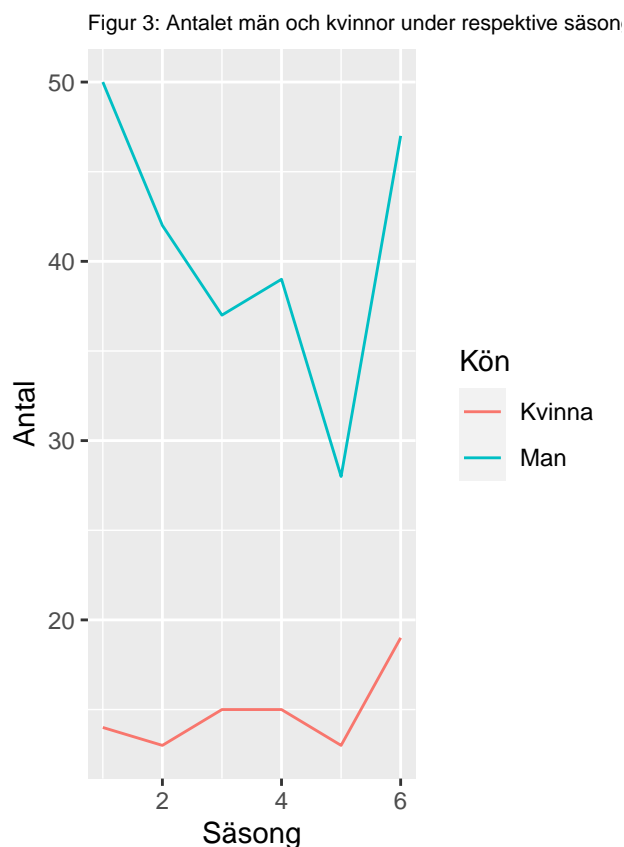
Utifrån “Figur 3” och “Figur 4” kan slutsatsen att det är betydligt mer vanligt med manliga karaktärer än kvinnliga. Det går dock att se en växande trend för kvinnliga karaktärer i tv-serien, medan i böckerna verkar antalet kvinnliga och manliga karaktärer svänga mer frekvent. Utifrån dessa figurer kan även slutsatsen att antalet karaktärer i tv-serien är betydligt färre än i böckerna.

Figur 1: Fördelning för antal säsonger



Figur 2: Fördelning för antal böcker





Vad är egentligen sannolikheten att man dör?

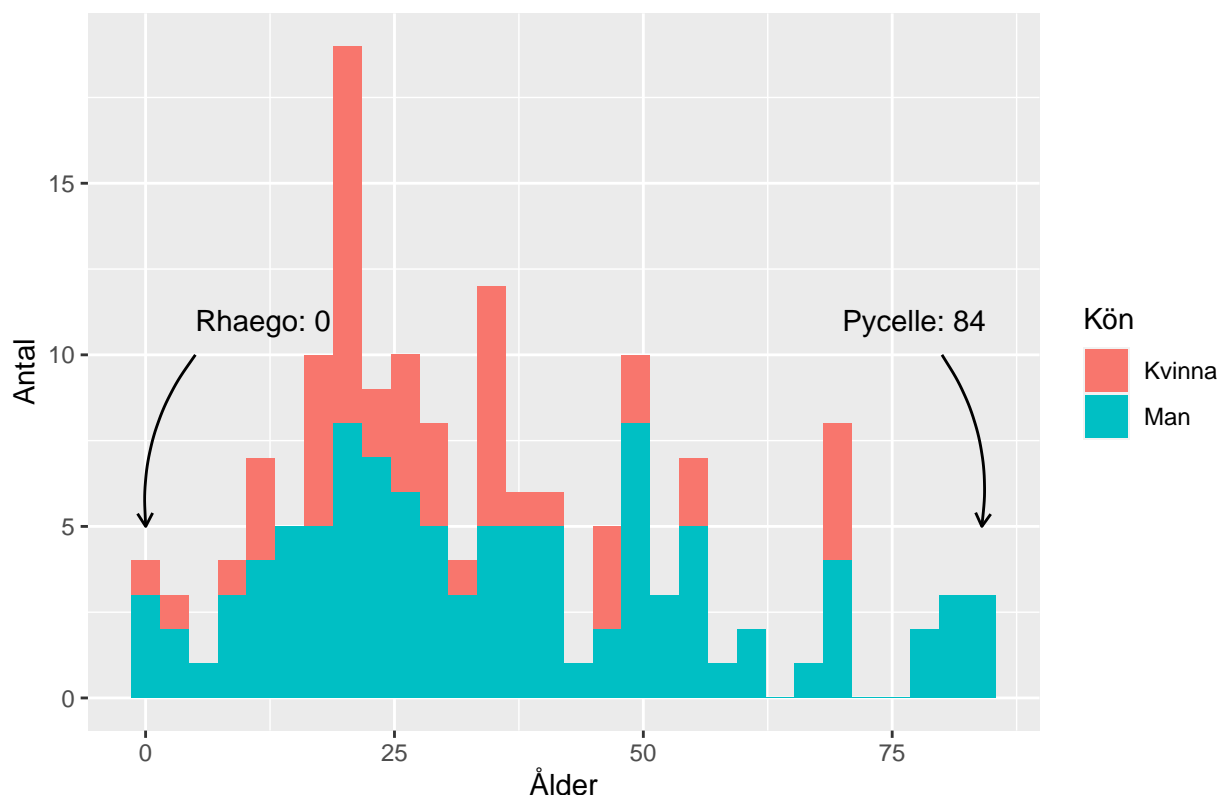
Det finns 1313 stycken unika karaktärer från alla dessa 10 böcker. I "Tabell 2" visas antalet som har dött, det vill säga utifrån alla dessa karaktärer har cirka 25% dött vilket motsvarar 331 stycken. Det går även att se att utifrån cirka 86.4% utav antalet döda var män och 14.6% var kvinnor. Detta kan ha varit en orsak till att kvinnor har haft en mer stabil kurva som kunde analyseras i "Figur 3" och "Figur 4". Det är värt att poängtera att denna information har tagits fram genom att endast tagit hänsyn till att information kring karaktärens död finns tillgänglig.

Genom att studera "Figur 5" kan slutsatsen att den förväntade åldern verkar ligga runt 25 år åldern, det finns dock vissa undantag för några karaktärer, vi ser till exempel att "Pycelle" är en utav de äldsta med en ålder på 84 år medan "Rhaego" inte hann fylla 1 år utan fick en livslängd på 0 år. Det kan även här vara värt att poängtera att dessa resultat har endast tagits fram utifrån de karaktärer som har haft både information angående födelsen samt deras död. Vi kan även utifrån denna graf tolka att livslängden verkar vara normalfördelad.

Table 2: Tabell 2: Andelen som har dött, totalt och för respektive kön

Totalt antal döda	Totalt procentuellt döda	Procentuellt döda Kvinnor	Procentuellt döda Män
331	0.2520944	0.1359517	0.8640483

Figur 5: Histogram över ålder



Vilket hus skapa en allians med?

Totalt finns det 292 olika hus, men vilket av dessa hus är egentligen bäst att bilda en allians med? I "Tabell 3" ser vi de 10 största husen som har flest medlemmar i Game of Thrones. Dock behöver inte alla inte alla medlemmar vara vid liv, i "Tabell 4" ser vi att det ändå är relativt många medlemmar som har dött, det lägsta antalet döda är "House Tyrell of Highgarden" med endast 7% som har dött. De med flest döda är "House Targaryen of King's Landing" med 55% döda. Om vi nu istället kollar på antalet levande medlemmar så är egentligen "House Tyrell of Highgarden" större än "House Targaryen of King's Landing". Det finns även massor av hus som har dött ut, dessa är dock ointressanta ty de endast bestod av ett fåtal medlemmar. I "Figur 6" ser vi ett histogram över de 4 största husen och ser att åldern för respektive hus är relativt jämt utspridd, dock kan det vara ett för få antal observationer för att få en bra uppfattning. Det kan eventuellt vara en normalfördelning som visar sig för åldern i "House Targaryen of King's Landing", men som tidigare nämnt det är för få antal observationer för att säga med säkerhet att det är en normalfördelning. Då det inte fanns någon information tillgänglig om åldern för medlemmarna i "House Tyrell of Highgarden" så kunde inte detta hus inkluderas i histogrammet.

Table 3: Tabell 3: Top 10 största husen

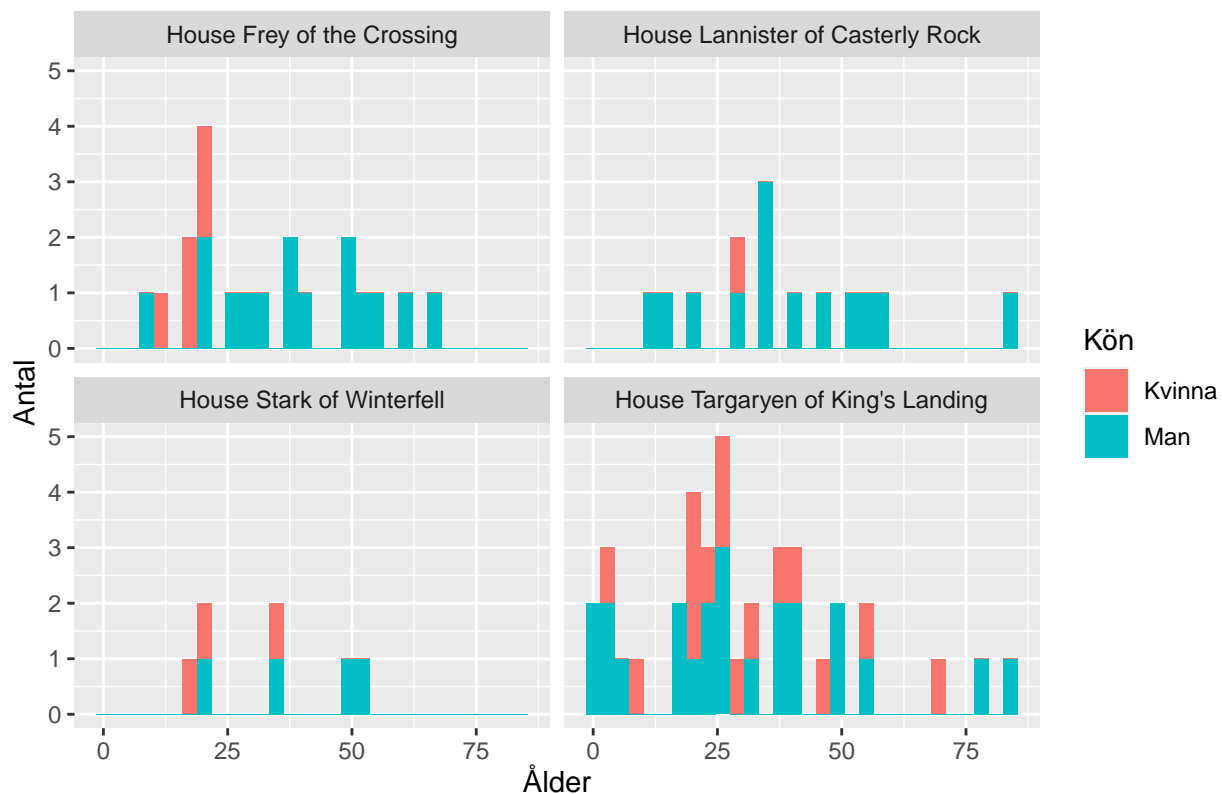
Hus	Antal medlemmar
House Frey of the Crossing	123
House Targaryen of King's Landing	84
House Stark of Winterfell	77
House Lannister of Casterly Rock	70
House Tyrell of Highgarden	43
House Greyjoy of Pyke	42
House Nymeros Martell of Sunspear	35

Hus	Antal medlemmar
House Baratheon of Dragonstone	27
House Arryn of the Eyrie	23
House Osgrey of Standfast	22

Table 4: Tabell 4: Top 5 största husen och respektive dödsantal

Hus	Totalt antal medlemmar	Antal levande	Antal döda	Procentuellt antal döda
House Frey of the Crossing	123	102	21	17%
House Targaryen of King's Landing	84	38	46	55%
House Stark of Winterfell	77	53	24	31%
House Lannister of Casterly Rock	70	43	27	39%
House Tyrell of Highgarden	43	40	3	7%

Figur 6: Histogram över åldern för de 4 största husen



Diskussion/Resultat

Då det egentligen inte finns något direkt resultat, då det inte finns någon tes utan detta projekt handlande endast om att förse läsaren med så mycket onödigt information som möjligt angående Game of Thrones. Därav kommer detta stycke vara en blandning av diskussion och resultat.

Om vi börjar med att diskutera API:n och hur informationen har tagits fram. Det optimalaste hade varit att endast tagit ut den sökta datamängden för varje moment, men på grund av att man måste kalla på API:n

flera gånger gör att det kommer ta tid för varje delmängd att plockas ut, samt att det finns ett tak på hur många gånger API:n får användas per dag. Detta gjorde det enkelt att bestämma att läsa in all data som fanns tillgänglig på API:n och sedan arbeta utifrån den. Jag har även valt att helt och hållet lita på att informationen från API:n stämmer till 100%, detta behöver inte vara fallet, som exempel vet jag "Cersei Lannister" är död, dock har inte detta tagits med i API:n, en orsak kan vara att hon eventuellt dog senare än säsong 6 eller bok nummer 10, detta är något jag har kollat upp utan endast reflekterat över.

Jag kan även utifrån vad som har analyserats ovan även tycka att det mest optimala för att få mest rampljus vore att vara en kvinna i tv-serien ty det är fler män som är karaktärer vilket gör att om tiden i rampluset skulle fördelas lika mellan män och kvinnor skulle kvinnor i snitt ha längre tid rampluset än män då det är betydligt färre. Vi kan även se att kurvan för antalet kvinnor inte lika "ostabil" i antalet män i böckerna och i serien. Denna teori styrks även om vi lägger den informationen med att antalet som har dött genom alla böcker, och bortser att böckerna och serien är två olika "världar" så är det 86% av de döda män så därav är det också bäst att vara en kvinna. Om jag hade varit en kvinna i Game of Thrones hade det enligt "Figur 5" sett ut som att jag hade haft en kort livslängd, då det kan analyseras att fler män har längre livslängd än kvinnorna, detta kan dock påverkas av att antalet observationer för kvinnor är betydligt lägre på grund av antalet kvinliga karaktärer är redan så få från början. Huset som jag sedan skulle vilja tillhöra är "House Tyrell of Highgarden", detta på grund av de har det lägsta dödsantalet, dock finns ingen information kring livslängden för detta hus men om jag istället utgår ifrån "Figur 5" som visar livslängden utan hänsyn till vilket hus karaktärerna tillhör så är optimalast. Om hus hade valt efter "Figur 6" så hade jag tagit "House Targaryen of King's Landing", även fast de har störst andel döda så har de mest information kring livslängden för kvinnorna vilket gör att jag vet vilken förväntad livslängd jag har. Detta kan dock diskuteras om man hellre vill veta sin förväntade livslängd eller inte, det vill säga vill man veta när man ska dö eller inte.

Om vi ska sammanfatta allt, det finns säkert ett mer effektivare sätt att samla in datamängden från API:n men enligt min egna åsikt tyckte jag att detta var lättast. Jag har valt att förlita mig till 100% att informationen som API:n generar är helt korrekt. Ett återkommande problem verkar även vara att observationerna verkar inte vara tillräckligt många för att kunna göra helt säkra slutsatser, till exempel är det väldigt få observationer kring kvinnors livslängd. Utifrån all information som har samlats in är det mest optimala för överleva och få mest rampljus i böckerna så som i serien är att vara en kvinna som tillhör huset "House Tyrell of Highgarden", detta på grund av att kvinnor har lägst andel döda i ett hus med de lägsta antalet döda.