

Olümpiamängude analüüs

Martin Sona

24.10.2019

1 Sissejuhatus

Töö eesmärgiks on analüüsida andmestiku, mis sisaldub endas informatsiooni kõigi inimeste kohta, kes on osalenud olümpiamängudel aastatel 1896-2016. Autor püstitas kolm hüpoteesi:

1. meeste ja naiste osakaal olümpiamängudel on ajapikku tasakaalustunud;
2. inimeste seas, kelle vanus on vahemikus 25-30, on kõige rohkem medalisaaajaid;
3. osalejate arv on kasvava trendiga.

Lisaks uurime olümpiamängude osalejate kaalu ja pikkust.

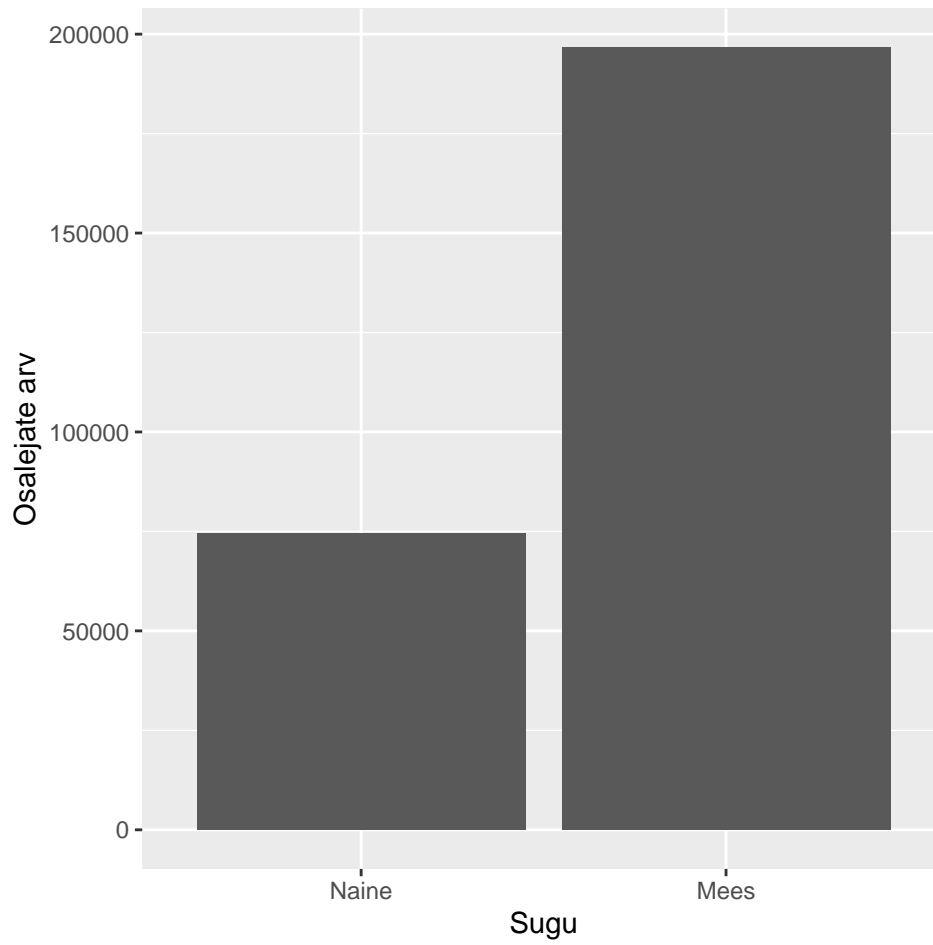
2 Andmestiku esmane kirjeldus

Andmestik koosneb 271116 reast ja 15 veerust. Andmestiku real on osaleja ning veergudes informatsioon selle osaleja kohta. Osaleja kohta on 15 erinevat tunnust: ID, nimi, sugu, vanus, pikkus, kaal, tiim, riik, mängud, aasta, aastaeg, linn, sport, spordiala, medal. Kokku on 10 faktorilist, 4 täisarvulist ja 1 reaalarvuline tunnus.

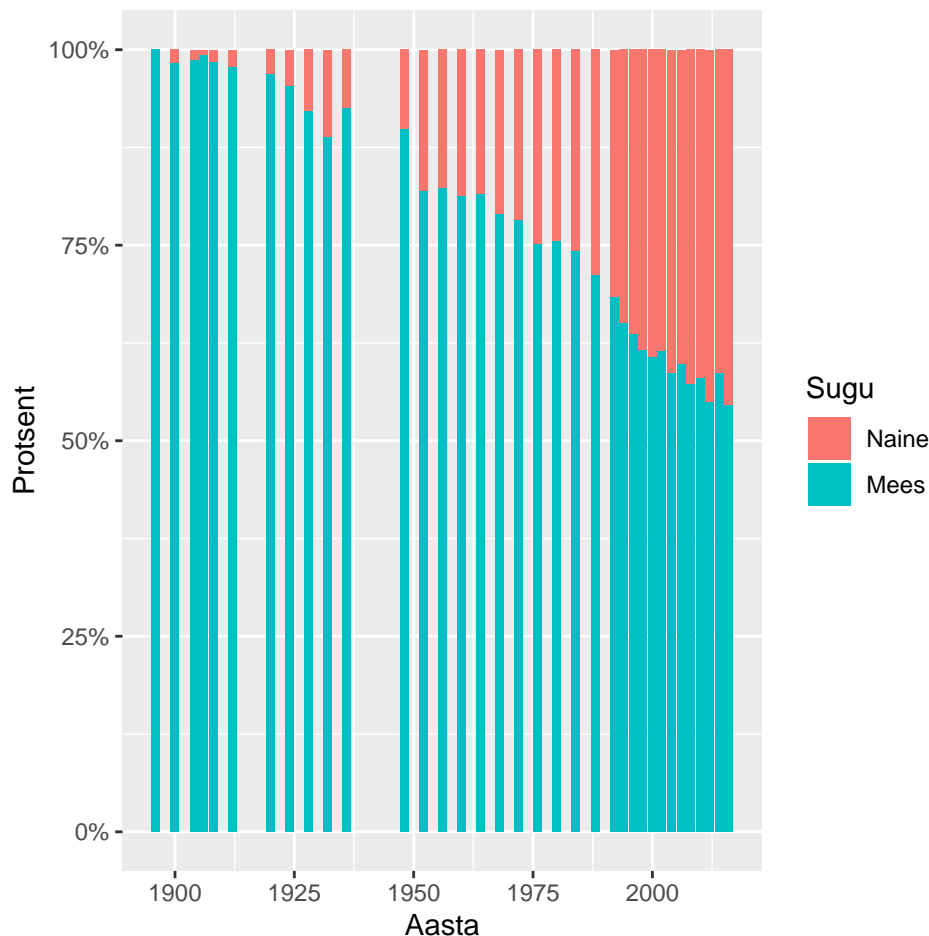
3 Analüüs

3.1 Olümpiamängude sooline jaotus

Esimesel joonisel on näha meeste ja naiste soolist jaotust olümpiamängudel. Siit on näha, et mehi on olümpiamängudel olnud peaaegu 3 korda rohkem kui naisi. Teisel joonisel on näha meeste ja naiste soolise jaotuse osakaalu läbi aastate, kus tuleb välja see, et mida aeg edasi, seda rohkem on sooline osakaal tasakaalustunud ning et viimasel olümpial oli naisi peaaegu sama palju kui mehi, seega meie esimene hüpotees on leidnud kinnitust.



Joonis 1: Naiste ja meeste jaotus.



Joonis 2: Naiste ja meeste osakaal läbi aastate.

3.2 Osalejate vanus, pikkus ja kaal

Lisaks soole on andmestikus ka informatsiooni osalejate vanuse, pikkuse ja kaalu kohta. Selleks, et neid uurida, teeme kõigepealt kaks tabelit, kus üks sisaldab informatsiooni kõigi osalejate kohta, teine aga medalisaajate kohta ning vaatleme nende karakteristiklike tunnuseid. Tabelitest on näha, et medali saajate ja osalejate vahel väga suuri erinevusi karakteristiklike tunnuste vahel ei ole, medalisaajad on keskmiselt natukene pikemad ja kaaluvad rohkem. Huvitav on aga see, et minimaalsed ja maksimaalsed väärtused erinevad nii drastiliselt. Lisaks, uurides neid tunnuseid, jääb mulje nagu kõik kolm on normaaljaotusega.

Table 1: Osalejate tunnused

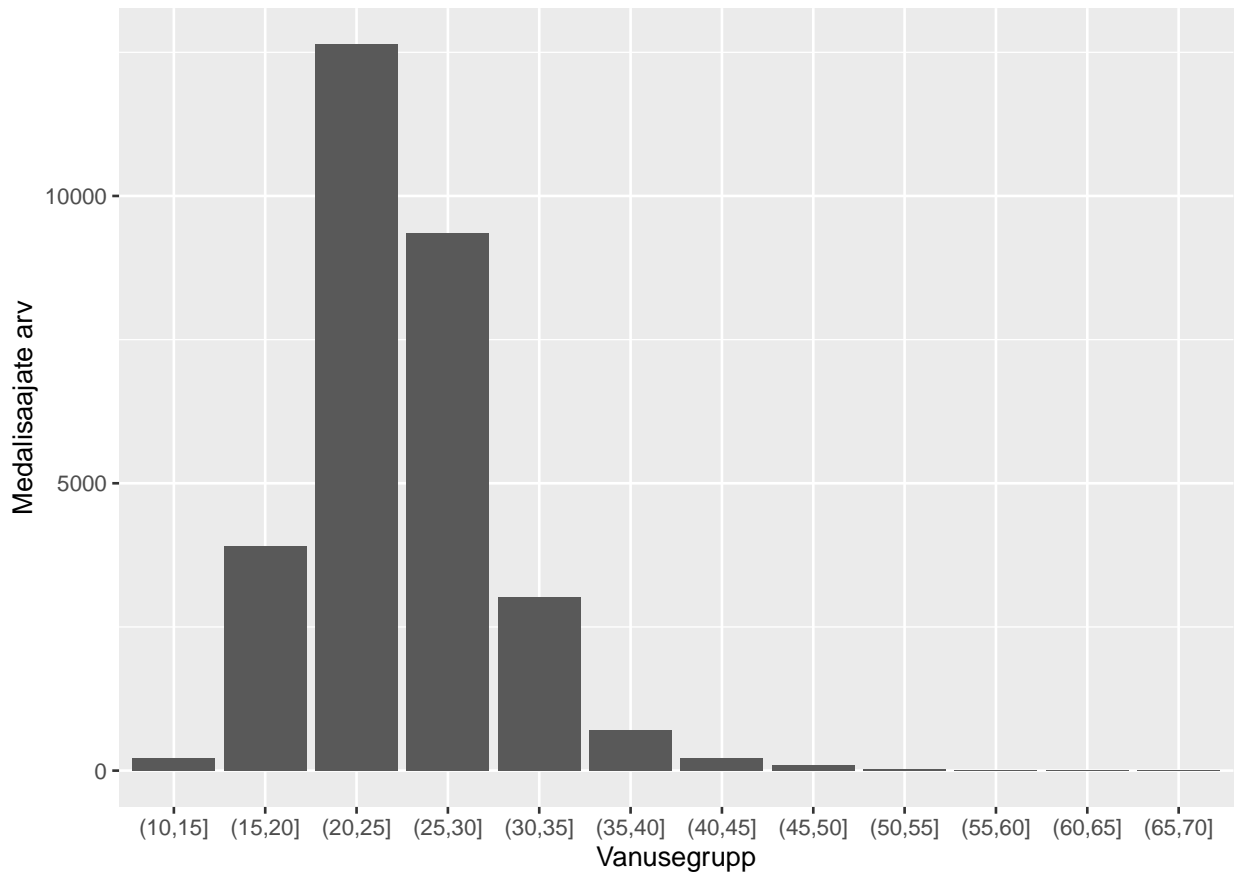
	Vanus	Pikkus	Kaal
	Min. :10.00	Min. :127.0	Min. : 25.0
	1st Qu.:21.00	1st Qu.:168.0	1st Qu.: 60.0
	Median :24.00	Median :175.0	Median : 70.0
	Mean :25.56	Mean :175.3	Mean : 70.7
	3rd Qu.:28.00	3rd Qu.:183.0	3rd Qu.: 79.0
	Max. :97.00	Max. :226.0	Max. :214.0
	NA's :9474	NA's :60171	NA's :62875

Table 2: Medali saajate tunnused

Vanus	Pikkus	Kaal
Min. :10.00	Min. :136.0	Min. : 28.00
1st Qu.:22.00	1st Qu.:170.0	1st Qu.: 63.00
Median :25.00	Median :178.0	Median : 73.00
Mean :25.93	Mean :177.6	Mean : 73.77
3rd Qu.:29.00	3rd Qu.:185.0	3rd Qu.: 83.00
Max. :73.00	Max. :223.0	Max. :182.00
NA's :732	NA's :8711	NA's :9327

3.2.1 Medalisaajate vanused

Teine hüpotees oli, et kõige rohkem on medalisaajaid vanuse vahemikus 25-30. Selle uurimiseks otsustasin teha histogrammi, kus ma jagasin vanused vahemikku pikkusega 5. Nagu jooniselt 3 on näha, siis hüpotees osutus valeks ning kõige rohkem oli medalisaajaid vahemikus 20-25.

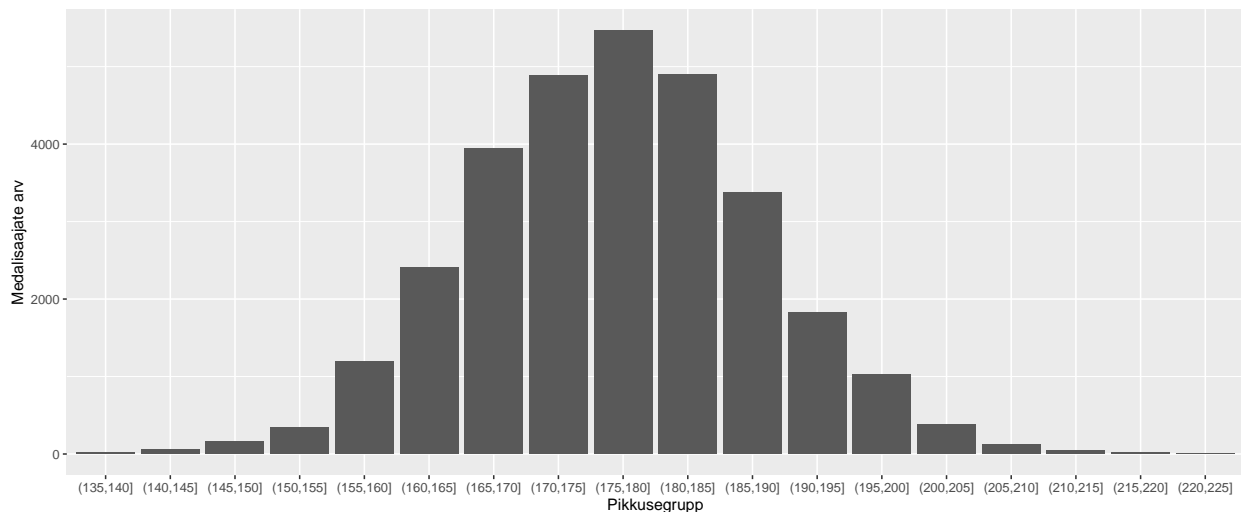


Joonis 3: Medalisaajate vanused.

3.2.2 Medalisaajate pikkused

Selleks, et uurida medalisaajate pikkusi otsustasin teha ka histogrammi, kus ma jagasin pikkused vahemikku pikkusega 10. Jooniselt 4 on näha, et olümpiamängudel osalejate pikkus on normaaljaotusega ning kõige

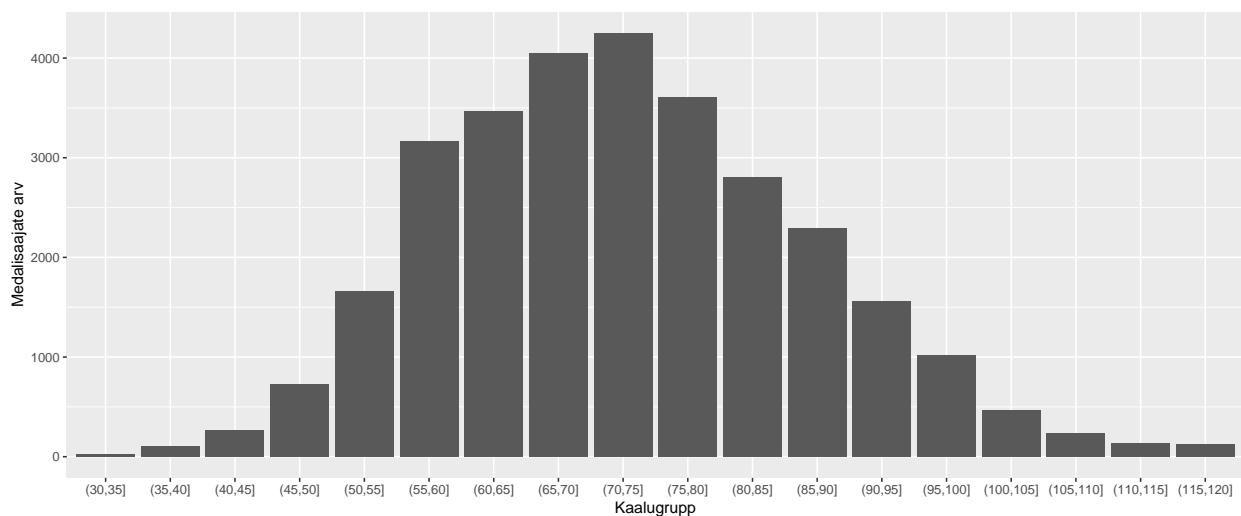
rohkem on pikkusega vahemikus 175-180 cm.



Joonis 4: Medalisajate pikkused.

3.2.3 Medalisajate kaalud

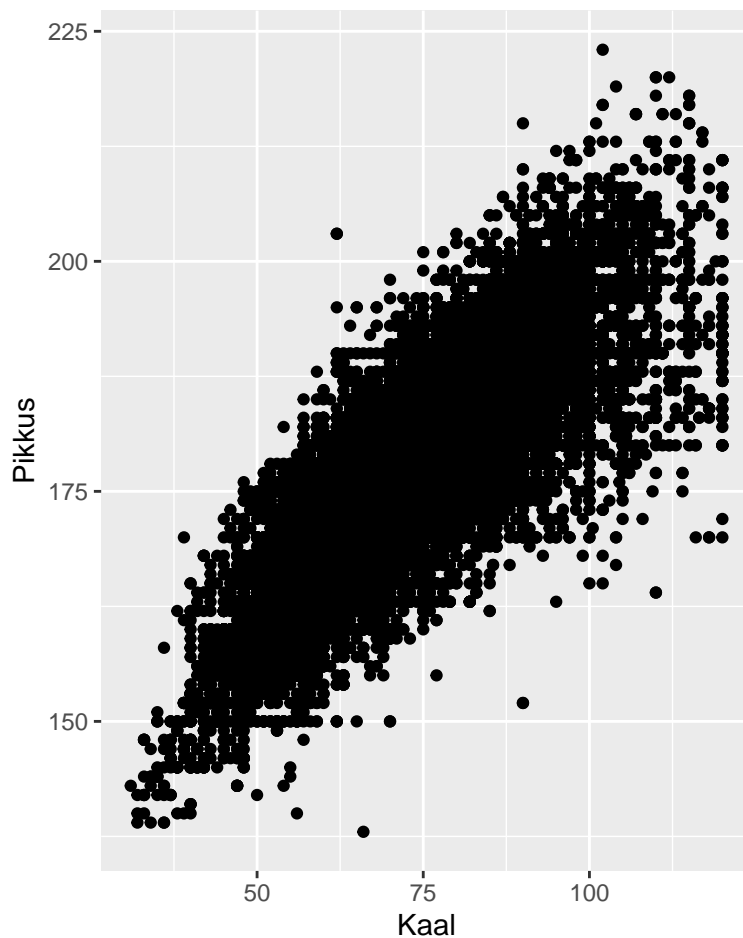
Kaalude puhul, erinevalt pikkustest, otsustasin mõned väärtused välja jätta, kuna erinevus minimaalse ja maksimaalse kaalu vahel oli liiga suur. Kuid jooniselt on näha, et kaalud on ka peaaegu normaaljaotusega ning kõige enam on medalisaajad kaaluga vahemikus 70-75 kg.



Joonis 5: Medalisajate kaalud.

3.2.4 Seos kaalu ja pikkuse vahel

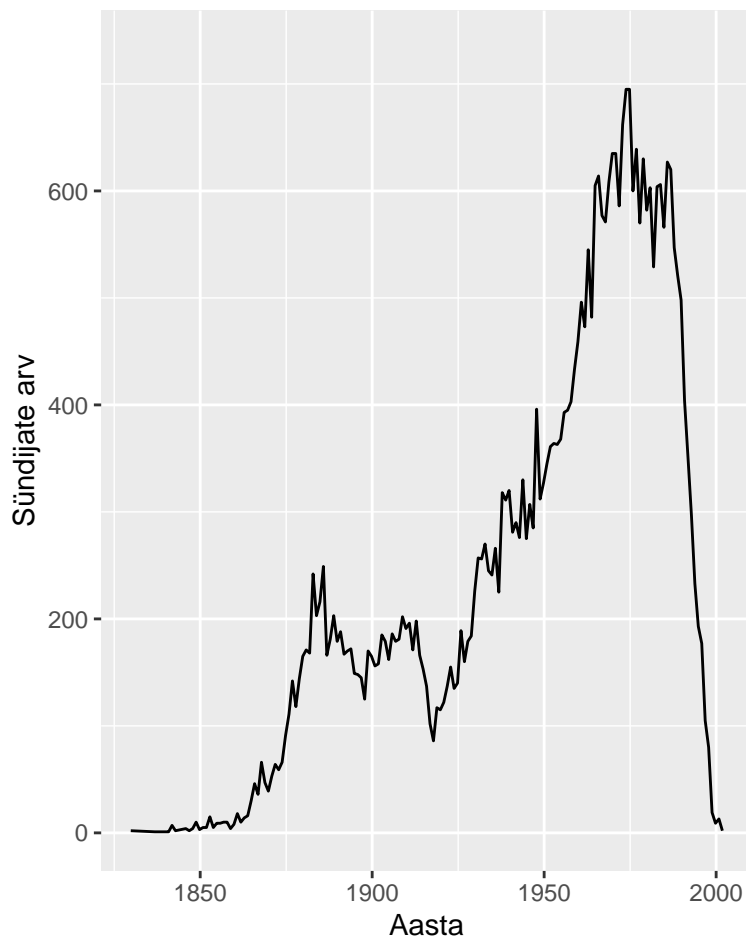
Vaatlesin ka seost kaalu ja pikkuse vahel. Jooniselt 6 tuli välja see, et suurt hajuvust väga ei ole, punktid paiknevad ilusasti tõusvas pilvekeses.



Joonis 6: Kaalu ja pikkuse vaheline seos.

3.3 Medali saajate sünniaastad

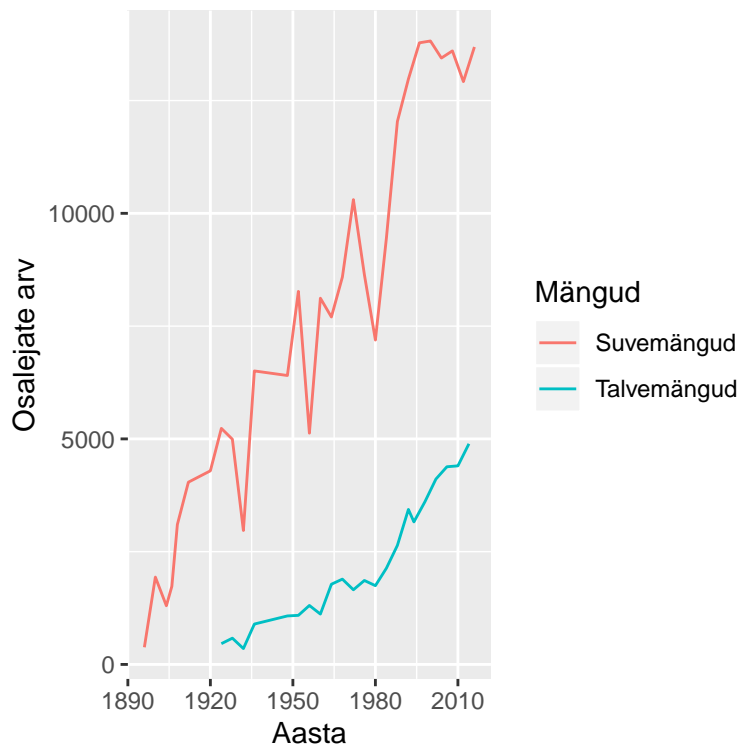
Otsustasin uurida ka medalisaajate sünniaastaid ning nende esinemiste sagedusi. Selleks tegin joondiagrammi. Jooniselt 7 tuli välja see, et medalisaajate sünniaastad on kasvava trendiga, mis ilmselt tähendab seda, et mida aeg edasi seda rohkem on tekkinud uusi spordialasid ning sellega on ka medalisaajate arv tõusnud.



Joonis 7: Medali saajate sünniaastad.

3.4 Osalejate arv olümpiamängudel.

Kolmas hüpotees oli, et osalejate arv olümpiamängudel on kasvava trendiga. Selle veendumiseks otsustasin teha joondiagrammi, kus üks joon on suvemängude kohta ning teine talvemängude kohta. Tegin kaks joont, kuna üldiselt suvemängudel on rohkem osalejaid kui talvemängudel, seega ei oleks saanud head ülevaadet, kui oleks ainult aastaid vaadanud. Joonis 8 on punane joon suvemängude kohta ning sinine talvemängude kohta. Jooniselt on näha, et mõlemad mängud on kasvavad, seega meie kolmas hüpotees leidis kinnitust.



Joonis 8: Osalejate arv olümpiamängudel läbi aastate.

4 Kokkuvõtte

Esimene ja kolmas hüpotees leidsid kinnitust, teine aga mitte. Tuli välja siis, et naiste ja meeste osakaal on tasakaalustumas; olümpia osalejate ja medalisaajate seas leidub igast vanusest, kaalust ja pikkusest inimesi; osalejad vanuse vahemikus 20-25 on kõige edukamad ning olümpial osalejate arv on tõusvas trendis. Lisaks, ma ei tea kui legitiimne muidugi see info on, siis isegi 73 aastasel on veel võimalik medal olümpialt endale saada.