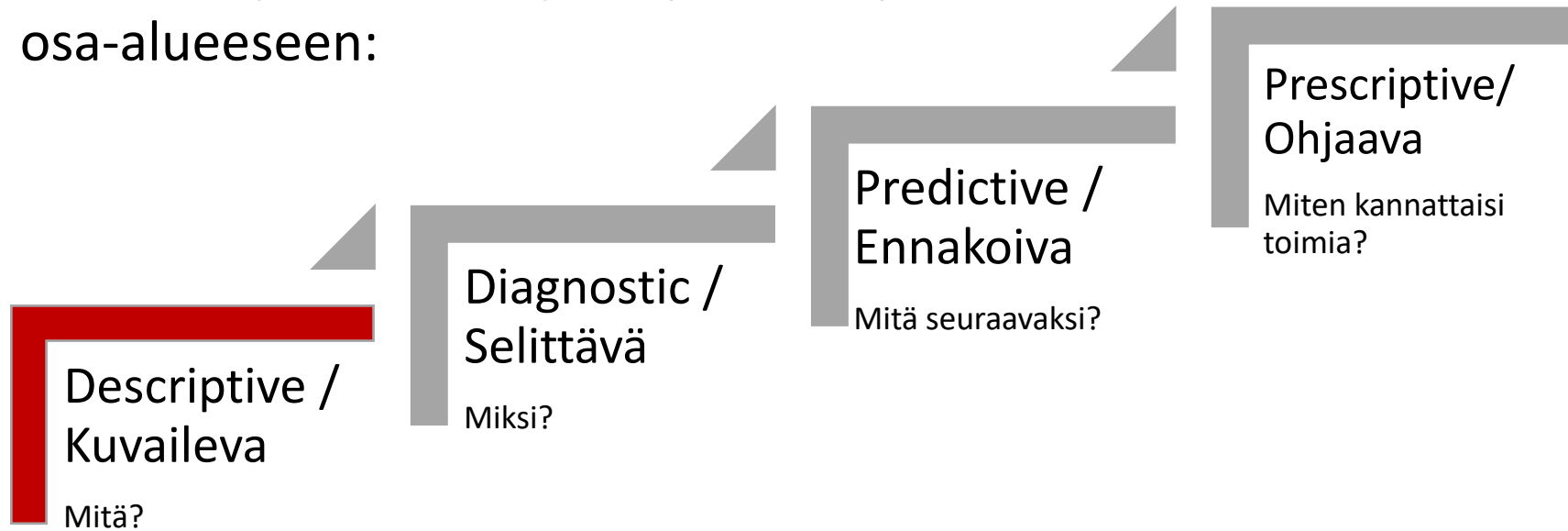


Kuvaileva analytiikka

Keskeiset menetelmät

Data-analytiikka

Data-analytiikka on tapana jakaa neljään osa-alueeseen:



Tämän esityksen aiheena on kuvaileva analytiikka

Kuvaileva analytiikka

Kategorinen muuttuja

- Frekvenssijakauma
 - Frekvenssitaulukko
 - Pylväskaavio (n-arvot näkyviin)

Huomaa! Jos lasket mielipideasteikon keskiarvon, niin käytä laskennassa vain asteikkoon kuuluvia arvoja.

Esimerkki. *1=Eri mieltä, 2=Jokseenkin eri mieltä, 3=Ei eri eikä samaa mieltä, 4=Jokseenkin samaa mieltä, 5=Samaa mieltä, 6=Ei kokemusta asiasta*

Tässä tapauksessa muuttujan arvoja 6 ei saa ottaa mukaan keskiarvon tai muidenkaan tilastollisten tunnuslukujen laskentaan.

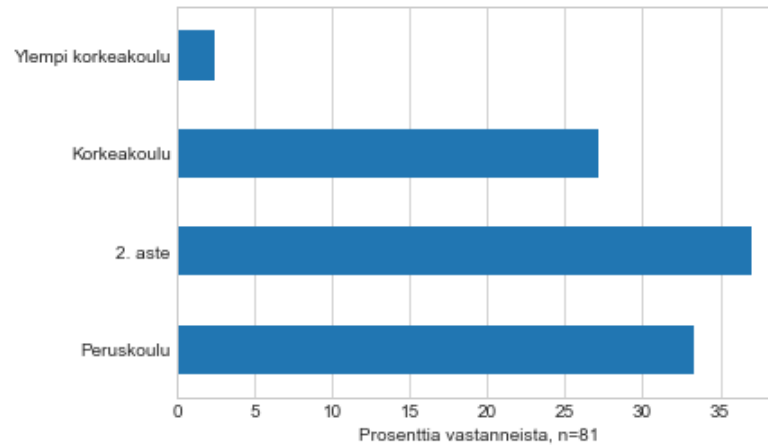
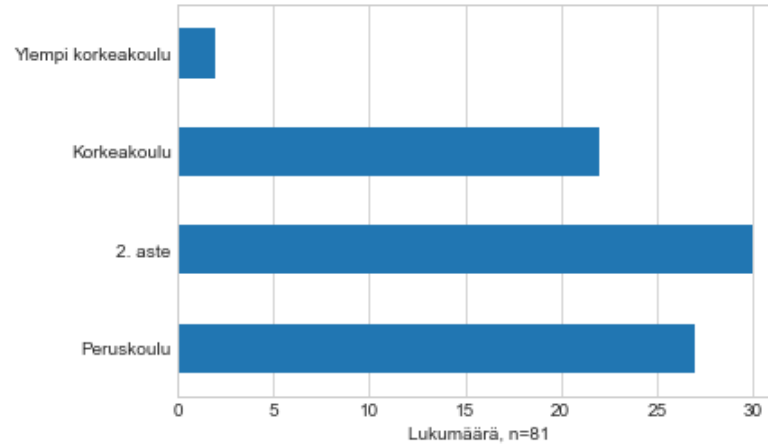
Määrällinen muuttuja

- Luokiteltu jakauma
 - Taulukko
 - Histogrammi
- Tilastollisia tunnuslukuja
 - Taulukko
 - Ruutu- ja janakaavio eli boxplot (n-arvot näkyviin)
 - Usean muuttujan keskiarvot pylväskaaviona (n-arvot näkyviin)

Kategorinen muuttuja

Frekvenssitaulukko ja pylväskaavio

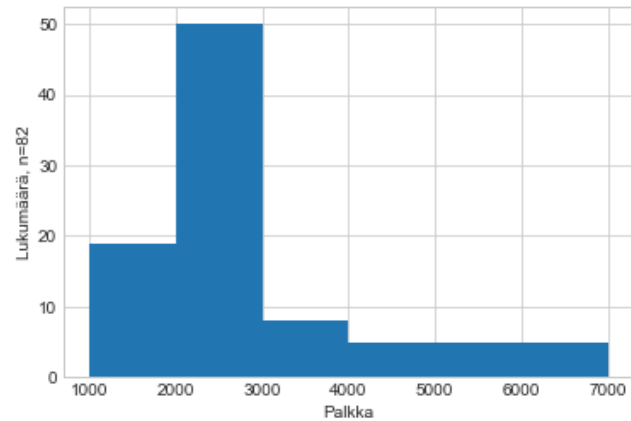
	Lukumäärä	%
Peruskoulu	27	33.3
2. aste	30	37.0
Korkeakoulu	22	27.2
Ylempi korkeakoulu	2	2.5
Yhteensä	81	100.0



Määrällinen muuttuja

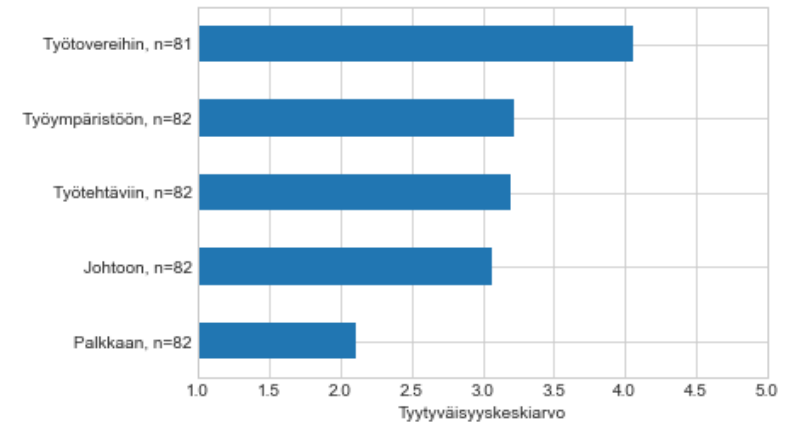
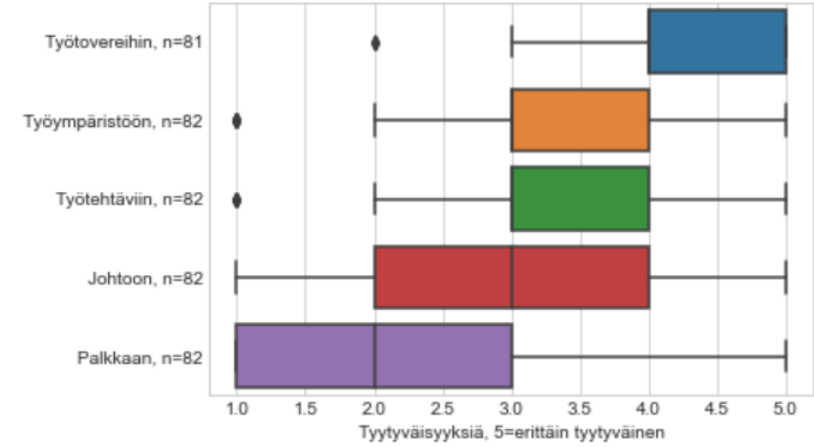
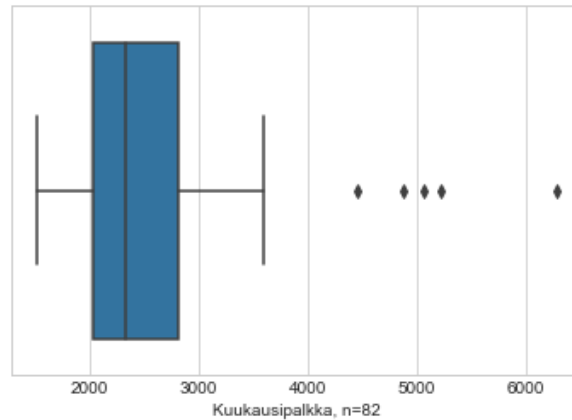
Luokiteltu jakauma

Palkkaluokka	Lukumäärä	%
(1000, 2000]	19	23.2 %
(2000, 3000]	50	61.0 %
(3000, 4000]	8	9.8 %
(4000, 7000]	5	6.1 %
Yhteensä	82	100.0 %



Tilastollisia tunnuslukuja

	Ikä	Palveluvuodet	Palkka
Lukumäärä	82.00	80.00	82.00
Keskiarvo	37.95	12.18	2563.88
Keskihajonta	9.77	8.81	849.35
Pienin	20.00	0.00	1521.00
Alaneljännes	31.00	3.75	2027.00
Mediaani	37.50	12.50	2320.00
Yläneljännes	44.00	18.25	2808.00
Suurin	61.00	36.00	6278.00



Aikasarja

- Viivakaavio havainnollistaa havaintojen ajallista kehittymistä.



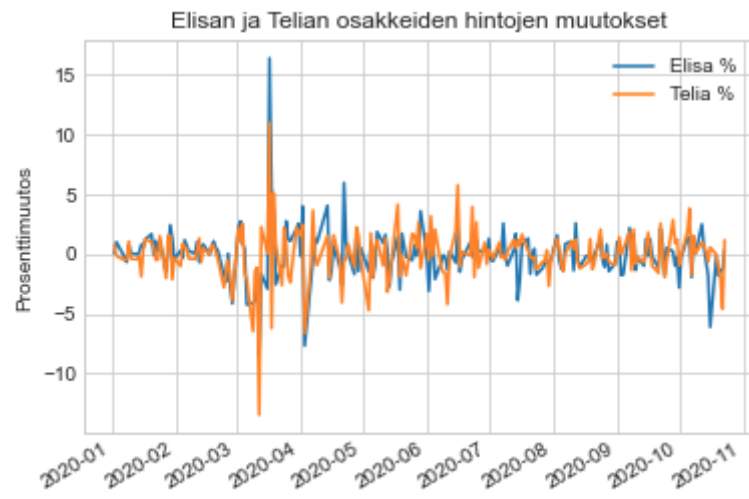
Aikasarjan vaihtelun tasoittaminen:

- Aggregointi: Havaintojen yhdistäminen alkuperäistä pidemmille aikaväleille, esimerkiksi päiväkohtaisista tiedoista kuukausikeskiarvoja.
- Liukuvat tunnusluvut: Esimerkiksi jokaiselle päivälle viimeisimpien 14 päivän keskiarvo.

Havaintojen sijasta prosenttimuutokset

- Prosenttimuutokset: Alkuperäisten havaintojen sijasta prosenttimuutos edellisestä ajankohdasta.
- Lyhyiden aikasarjojen kohdalla prosenttimuutoksia toisinaan havainnollistetaan viivakaavion sijasta pylväskaaviolla.

Aikasarjan kuvailua



Välineet kuvailevaan analytiikkaan

Kuvailevaa analytiikkaa voi harjoittaa monenlaisilla välineillä:

- **Taulukkolaskennalla** (esimerkiksi Excel)
 - Aikasarjojen analysointiin puutteelliset työkalut
- **Tilasto-ohjelmalla** (esimerkiksi SPSS)
- **Koodaamalla** (Python tai R)
 - Paras vaihtoehto, jos analysoidaan toistuvasti rakenteeltaan samanlaisia dataa tai jos analysoitava data päivittyy aika ajoin
 - Erinomaiset välineet aikasarjojen analysointiin
 - Tämän esityksen kaaviot ja taulukot laadittu Pythonilla