|  |
| --- |
| **Námskeiðstitill:** |
| **Stutt lýsing**: Hvað eru gagnavísindi (e. Data science) og af hverju eru þau svona vinsæl núna? Á námskeiðinu verður farið ítarlega í gegnum mikilvægustu þætti gagnavísinda. Nemendur munu læra að nota R (RStudio) og gagnavísindi m.a. til að takast á við verkefni sem erfitt er að fást við í töflureikni, læra *grammar of graphics* til að búa til gröf, einfalda ferlið frá hráum gögnum að lokaafurð og sjálfvirknivæða endurtekin verkefni. |
| **Lengri lýsing:** Á námskeiðinu verður farið ítarlega yfir notkun R (RStudio), einu vinsælasta forritinu þegar kemur að gagnavísindum. Nemendur læra að búa til skýrslur og glærur inni í R og þannig einfalda ferlið frá hráum gögnum að lokaafurð.  Farið verður yfir hvernig má hlaða inn gögnum úr mismunandi áttum. Þá verð varið yfir mikilvægustu atriðin við að hreinsa gögnin á hagkvæman hátt og koma þeim á *tidy* snið. Lögð verður rík áhersla á að nemendur öðlist góða færni í að draga fram mikið af upplýsingum úr gögnum, stórum sem smáum.  Þá verður farið ítarlega yfir hvernig má nota *grammar of graphics* til að búa til glæsileg gröf. |
| **Á námskeiðinu er fjallað um:** Áherslupunktar – mega vera fleiri eða færri |
| * Inngang að R   + Búa til verkefni (e. Project) með RStudio   + RStudio umhverfið   + Útreikningar með R   + Búa til eigin föll   + Lesa inn gögn úr csv, Excel, af netinu, úr gagnagrunnum og úr pökkum líkt og Eurostat og OECD |
| * RMarkdown   + Skýrslur (PDF, Word og HTML)   + Glærur (PDF, Power Point og HTML)   + Nota R kóða í skýrslum og glærum |
| * Gagnavinnsla með dplyr pakkanum   + Velja athuganir sem uppfylla þau skilyrði sem við setjum   + Velja breytur sem við viljum nota   + Búa til nýjar breytur   + Búa til tölfræðisamantekt   + Framkvæma útreikninga niður á hópa   + Skeyta saman gagnasettum |
| * Myndræn framsetning gagna með ggplot2 pakkanum   + Mismunandi tegundir af gröfum (scatter, histogram, linegraph, boxplot o.fl)   + Breyta litum, stærð og lögun   + Nota *breytu* til að stjórna litum, stærð og lögun   + Búa til fylkja-gröf (mörg lítil gröf sett saman í eitt stórt graf)   + Vista gröf   + Breyta kyrrstæðum gröfum í lifandi gröf |
| * Búa til *tidy data* með tidyr pakkanum   + Breyta gögnum úr *long format* í *wide format* og öfugt   + Skipta breytu/dálk í margar breytur/dálka   + Skeyta saman dálkum/breytum * Fylla inn í *missing values* |
| * Nota stýrisetningar (if-else), *for loop-*u, kynning á functional programming og tímasetning R forrita   + Læra að nota *for loop*-u   + Læra að nota if-else setningar og case\_when   + Læra að nota helstu föllin úr purrr pakkanum   + Læra að nota taskscheduleR pakkann til að tímasetja R forrit |
| **Ávinningur þinn:** Áherslupunktar – mega vera fleiri eða færri |
| * Stóraukin hæfni og hraði í greiningu og vinnslu gagna |
| * Framkvæma útreikninga og búa til gröf sem erfitt eða ómögulegt er að gera t.d. í hefðbundnum töflureikni |
| * Láta tölvuna vinna einhæf verkefni fyrir þig |
| * Fækka algengum villum með því að halda greiningu og skrifum / gerð kynninga á sama stað. |
| **Fyrir hverja:** |
| Námskeiðið er fyrir þá sem unnið hafa eitthvað með gögn og vilja stórbæta og útvíkka þekkingu sína í greiningu gagna. |
| **Aðrar upplýsingar:** |
| Nemendur þurfa að koma með eigin tölvu. Ætlast er til að nemendur setji upp R og RStudio fyrir fyrstu kennslustund. Nemendur fá ítarlegar leiðbeiningar í pósti hvernig það er gert. Engin ein kennslubók er notuð beint í þessum áfanga. Áhugasömum nemendum sem vilja ná góðum tökum á gagnavísindum er bent á eftirfarandi bækur:   * R for Everyone: Advanced Analytics and Graphics eftir Jared P. Lander * R for Data Science eftir Garrett Grolemund og Hadley Wickham * Aðgengileg frá höfundum hér: <https://r4ds.had.co.nz> |
|  |
|  |

Námskeiðslýsingin er birt í ofangreindri röð á vefsíðu EHÍ.

Fyrir ofan lýsinguna er birtur *skráningarfrestur, nafn kennara, hvenær, hvar* og *verð*. [www.endurmenntun.is](http://www.endurmenntun.is/)