

NEKAJ MOŽNIH NAMENOV

Izboljšati trenutno predstavitev podatkov.

Pokazati napredek v medicini, ki izvira bodisi iz boljšega zdravljenja ali pa hitrejše diagnostike → Fokus na nekaj vrst raka.

Predstaviti ključne informacije za splošno laično publiko (*angl.* cancer factsheet), npr. uporabiti mobilno aplikacijo kot način dostopa do laične publike. → Izbor glavnih kazalnikov.

OMEJITVE

Preživetje po regijah ni objavljeno.

Dejavnikov tveganja ne vključujemo v prikaz, saj zanje ni vir podatkov Slora, ampak bi se morali nasloniti na študije teh dejavnikov, ki pa jih je več kot dvajset. To je že področje analitske epidemiologije, medtem ko so podatki iz Slore namenjeni opisni epidemiologiji.

ČASOVNE VRSTE

Za prikaz je smiselno uporabiti vse razpoložljive točke oz. vsaj kakih 20 točk, da se razbere trend.

Na ravni razčlenjenosti, ki je uporabljena v bazi Slora, ni omejitev glede časovne primerljivosti (npr. zaradi spremembe klasifikacij).

STAROSTNI RAZREDI

Podatki so na voljo po petletnih starostnih razredih, kar je npr. primerno za prikaz porazdelitve po starosti (histogram, poligon).

Če je potrebno uporabiti manjše število kategorij, potem je smiselno oblikovati naslednje starostne skupine:

0 – 19
20 – 34
35 – 49
50 – 64
65 – 75
75 +

Če bi bilo tudi šest kategorij preveč (npr. preveč linij v linijskem grafikonu), je smiselno združiti prva dva razreda (0 – 34) ali celo prve tri razrede (0 – 49). Rak je namreč najbolj prisoten pri starejših.

Posamezne starostne skupine bi bile zanimive samo za specifične namene, npr. starostna skupina 20 – 64 bi lahko bila zanimiva z vidika vračanja na delovno mesto.

Z vidika ocenjevanja kakovosti zdravljenja je morda smiselno izključiti skupino 80+, ker ima veliko oseb omejitve pri izbiri terapij.

VRSTE RAKA / LOKACIJA RAKA

V bazi je veliko vrst raka, za večino prikazov nima smisla prikazati čisto vseh vrst. Kako narediti izbor?

- Najpogostejše vrste raka, npr. Top5 + Ostali, Top9 + Ostali, Top10 + Ostali.
- Raki, s katerimi so povezane kake aktivnosti, npr. presejalni programi (ZORA, DORA, SVIT). Pri raku debelega črevesja in danke se vidi naraščajoči trend do 2010, nato pa upad kot rezultat uvedbe presejalnega programa. Pri raku ščitnice velja opozorilo, da so se številke za incidenco povečale, odkar se upoštevajo mejno maligni raki, od katerih osebe ne umrejo, breme ni večje, tudi po spolu ni razlike.

SPOL, STADIJ

Obe spremenljivki sta zelo zanimivi za prikaz.

STOLPČNI GRAFIKONI

Primerjava incidence in prevalence ne gre na isti grafikon, ker so velikostni redi pojava zelo različni (prevalenca precej večja od incidence).

Ležeči grafikon ima to prednost pred pokončnim, da omogoča daljše opise posameznih stolpcev/kategorij.

Razvrstitev, ki sledi razvrstitvi v klasifikaciji, ne bi bila konsistentna, čeprav naj bi si lokacije sledile od glave navzdol. Idealen vrstni red stolpcev je po velikosti, kar pa zahteva dinamično razvrščanje, če hočemo prikazati razvrstitev skozi čas. Primera, kako bi to lahko izgledalo:

<https://public.flourish.studio/visualisation/270258/>

https://www.reddit.com/r/dataisbeautiful/comments/fz2lod/oc_coronavirus_deaths_per_capita_by_country/

LINIJSKI GRAFIKONI

Linijski grafikon lahko hitro postanejo 'spaghetti'. Ena rešitev je animacija, npr.

<https://www.idc.com/promo/smartphone-market-share/vendor>. Druga rešitev so small multiples, kot smo jih srečali na predavanjih. Tretja rešitev je kompromisna in se osredotoča samo na začetek in konec opazovanega obdobja (<https://ourworldindata.org/cancer-death-rates-are-falling-five-year-survival-rates-are-rising>).

ZEMLJEVIDI

Na ravni regij načeloma ni smiselno prikazovati absolutnih podatkov, saj se število prebivalcev po regijah zelo razlikuje. Smiselno se je torej osredotočiti na stopnje.

Kako opredeliti skalo in razrede:

- relativno v vsakem letu (npr. razdelitev na pet kvintilnih razredov → poudarek na vsakokratni geografski primerjavi) ali
- absolutno (npr. razdelitev variacijskega razmika na določeno število razredov → poudarek na medčasovni primerjav)

Pri relativni opredelitvi razredov je največji problem pravilno skomunicirati glavno sporočilo (npr. nadpovprečno in podpovprečno incidenčno/prevalenčno stopnjo v posameznem letu) in preprečiti direktne medčasovne primerjave (povečanje/zmanjšanje stopnje iz leta v leto).

Pri absolutni opredelitvi razredov se nam pri prikazu razvoja raka skozi čas lahko zgodi, da bodo vse regije na začetku obdobja na eni skrajnosti skale, na koncu obdobja pa na drugi skrajnosti skale (npr. če je ob rasti pojavnosti raka skozi čas časovna razlika precej večja od medregijskih). Ena rešitev bi bila v večjem številu razredov, vendar pa s tem trčimo ob drugo omejitev oz. priporočilo, naj bi razredov do sedem, da bomo še dobro zaznali posamezno barvo.

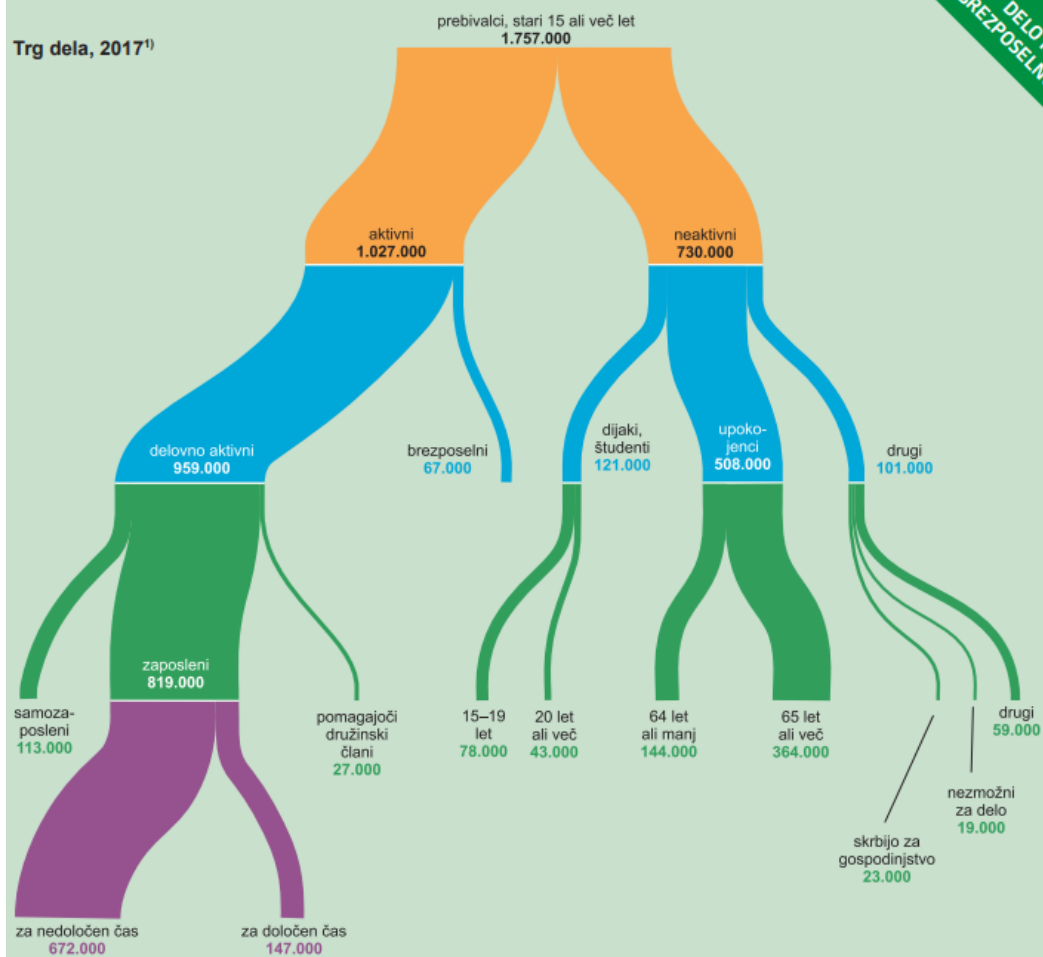
Izbira barv za razrede je zelo zahtevna. Priporočam Colorbrewer2.0 (<https://colorbrewer2.org/>). Na splošno o uporabi barv lahko priporočim blog Lise Charlotte Rost (<https://blog.datawrapper.de/colors/>), ima pa tudi članek o orodjih, če vam Colorbrewer2.0 ne bi zadoščal (<https://blog.datawrapper.de/colorguide/>).

STRUKTURE

Glede na omejitve strukturnih krogov in za potrebe kompleksnejše vizualizacije je na voljo kar nekaj rešitev. Za podatke, ki se lahko členijo na več ravneh, je uporaben tudi Sankeyjev diagram, ki je v osnovi namenjen prikazu tokov, npr.

Trg dela, 2017¹⁾

DELO IN
BREZPOSELNOST



1) Nekateri seštevki se zaradi zaokroževanja ne ujemajo.
Vir: SURS