

Upoštevanje prestopnih let pri tedenskih medletnih primerjavah

mz

09.December 2024

1 Intro

Pri medletnih tedenskih primerjavah podatkov, ki jih trenutno izračunavamo za podatke iz davčnih blagajn, je treba v praksi rešiti problem prestopnih let. Ta dokument opisuje uporabljeno rešitev.

2 Problem

Teden, ki v prestopnem letu vsebuje 29.2., ima v prejšnjem letu samo 6 ujemajočih ali ekvivalentnih datumov. Teden v po-prestopnem letu, ki se začne z datumi od 23. do 1.3. ima v prejšnjem letu en “dodaten” dan za primerjavo, ki ga izgubimo, če primerjamo samo ekvivalentne datume.

Čiste rešitve ni, iščemo torej rešitev, ki:

- zagotovi, da se v obeh letih vedno primerja vseh 7 dni v tednu, kar je sploh pomembno pri podatkih kot so davčne blagajne, ki imajo izrazito tedensko nihanje;
- zagotovi, da se uporabi vse podatke, torej se ničesar ne zavrže popolnoma;
- je generična in relativno enostavna za razumeti, če že ne za implementirati..

3 Rešitev

3.1 Prestopno leto

V prestopnem letu teden, ki vsebuje 29.2., primerjamo s šestimi ekvivalentnimi datumi iz prejšnjega leta, ki jim kot sedmi dan dodamo povprečje dneva pred tem in dneva po tem.

Kot primer je spodaj prikazana situacija leta 2024:

- v letu 2023 vzamemo ekvivalentne datume od 26.2. do 3.3. (med katerimi manjka sobota)
- vzamemo povprečje obeh sobot na vsaki strani: 25.2. in 4.3. in jih dodamo ostalim šestim dnevom.

2024		2023		
ned	25.2.	sob	25.2.	×0.5
pon	26.2.	ned	26.2.	×1
tor	27.2.	pon	27.2.	×1
sre	28.2.	tor	28.2.	×1
čet	29.2.		—	
pet	1.3.	sre	1.3.	×1
sob	2.3.	čet	2.3.	×1
ned	3.3.	pet	3.3.	×1
pon	4.3.	sob	4.3.	×0.5

Table 1: Primer izračuna v prestopnem letu

To pomeni, da so zneski za ti dve soboti v medletni primerjavi upoštevani 1.5-krat: enkrat v običajni primerjavi kot del normalnega tedna in pol-krat v primerjavi kot fiktivna sobota, ki nam sicer manjka.

3.2 Po-prestopno leto

V po-prestopnem letu imamo dve možnosti:

- ali se v tem letu teden začne s 1.3. in je torej lanski 29.2. pred tem tednom
- ali se v tem letu teden začne z 23. - 28.2. in je torej lanski 29.2. nekje na sredini tega tedna.

3.2.1 Teden se začne s 1.3.

Vrednost 29.2. povprečimo z naslednjim istim dnevom.

Kot primer je prikazana situacija v letu 2021:

- 29.2. pade na soboto, zato vzamemo povprečje 29.2. in sobote 7.3.
- ostalih šest dni se primerja z ekvivalentnimi datumi v letu 2020

	2021		2020	
	—	sob	29.2.	×0.5
pon	1.3.	ned	1.3.	×1
tor	2.3.	pon	2.3.	×1
sre	3.3.	tor	3.3.	×1
čet	4.3.	sre	4.3.	×1
pet	5.3.	čet	5.3.	×1
sob	6.3.	pet	6.3.	×1
ned	7.3.	sob	7.3.	×0.5

Table 2: Primer izračuna v po-prestopnem letu: teden se začne z 1.3.

To pomeni, da so zneski za ti dve soboti v medletni primerjavi upoštevani samo 0.5-krat, saj noben od teh datumov v letu 2020 ni prisoten v nobeni drugi tedenski primerjavi.

3.2.2 Teden se začne z 23. do 28.2.

Povprečimo vrednosti prvega in zadnjega ekvivalentnega datuma, ki vedno padeta na isti dan.

Kot primer je prikazana situacija v letu 2025:

- prvi in zadnji ekvivalentni datum sta 24.2. in 2.3., oboje soboti, zato vzamemo njuno povprečje
- V implementaciji se to povprečje prestavi na zadnjo od obeh sobot, medtem ko se vrednost 29.2. prestavi na prvo od obeh sobot, ampak to niti ni relevantno, ker se itak potem agregira na tedenski ravni, tako da je važno samo, da imamo 7 različnih dni za primerjavo.

2025		2024		
pon	24.3.	sob	24.3.	×0.5
tor	25.3.	ned	25.3.	×1
sre	26.3.	pon	26.3.	×1
čet	27.3.	tor	27.3.	×1
pet	28.3.	sre	28.3.	×1
	—	čet	29.3.	×1
pet	1.3.	pet	1.3.	×1
sob	2.3.	sob	2.3.	×0.5

Table 3: Primer izračuna v po-prestopnem letu: teden se začne z 23. do 28.2.

To spet pomeni, da so zneski za ti dve soboti v medletni primerjavi upoštevani samo 0.5-krat.

4 Implementacija

Trenutno je rešitev implementirana v grafu v poročilu indikatorjev za katerega koda se nahaja pri Maji v mapi `M:\analysis\umar_master_report\docs\children\` v fajlu `html_94_davcne.R`.