**Gagnavinnsla – Hvernig voru gögnin meðhöndluð?**

**Hreinsun á gögnum**

Þegar að gögnin voru komin yfir í R fór frekari skoðun í gang á töflunni timtaka\_data, sást þá að á sumum stöðum voru tómar raðir sem innihéldu aðeins id og hlaup\_id. Þeim röðum var hreinsað út og gert var ný tafla sem heitir timataka\_data2. Með hreinan gagnagrunn sem hægt væri að nota til frekari úrvinnslu á gögnunum hófst vinna við val á verkefni.

**Val og birting á efni og gröfum**

Til að finna áhugaverða einstaklinga til að fjalla um var skoðað hvaða einstaklingar voru oftast að keppa. Til þess að finna það var notað *count* til að telja hvað allir einstaklingar voru að keppa oft og síðan aðeins prentað út og geymt í sér töflu þá einstaklinga sem komu oftar en 14 sinnum fram. Aðeins átta einstaklingar voru oftar en 14 sinnum og af þeim voru þau Bergdís Eva Sveinsdóttir og Eyjólfur Guðgeirsson sem kepptu oftast eða 17 sinnum. Rúnar Sigurðsson keppti einnig 17 sinnum en ákváðum við að hafa aðeins þau Bergdísi og Eyjólf eftir að skoðað var hvar þau enduðu í keppnum hjá sér, Rank, sem sagt að þau enduðu hærra heldur en Rúnar.

Eftir að búið var að ákveða að hafa þau Bergdísi og Eyjólf, þá var farið í það að búa til sér gagnatöflur fyrir þau bæði eða bergdis\_data og eyjolfur\_data sem innihélt aðeins upplýsingar um þeirra keppnir úr timataka\_data2.

Til að geta borið þau mögulega saman þá þarf að finna hvort þau hafa keppt í sömu keppnum. Eftir að hafa keyrt þann kóða kom í ljós að þau höfðu aðeins keppt þrisvar sinnum í sömu keppnum. (Sjá mynd fyrir neðan)

A screenshot of a computer

Description automatically generated

Næsta skref var að gera línurit til þess að bera saman tíma þeirra í þessum hlaupum en einnig Rank s.s. hvar þau enduðu bæði tvö. Einnig vildum við sjá hvernig þau sem einstaklingar voru að ganga og þarf þá skoða þau hlaup eða hjólreiðar sem þau hafa tekið þátt í nokkrum sinnum. Við yfirlestur á bergdis\_data og eyjolfur\_data, þeirra eigin gagantöflum, sást að Bergdís hefði aðeins tekið þátt í Íslandsmótinu í Criterium oftar en tvisvar. Eyjólfur hefur einnig keppt oftar en tvisvar í Criterium sem og Fellahringnum. Með þær upplýsingar fyrir hendi var síðan gert þrjú línurit til að sýna árangur þeirra beggja í þessum keppnum milli ára með tilliti til tíma. Tvö fyrir báðar keppnir hans Eyjólfs og eitt fyrir hana Bergdísi.

Það næsta sem við fórum að skoða var hversu oft þau Bergdís og Eyjólfur voru að keppa ár hvert. Til þess var notað *count* til að einfaldlega telja þann fjölda af keppnum hjá þeim báðum og það sett upp í stöplarit. Áður en það var hægt þurfti að draga úr nafna dálknum yfir keppninar ártölin og búa til nýjan dálk. Með árin komin var síðan sameinað bergdis\_data og eyjolfur\_data (þeirra eigin gagnatöflur) til þess að hafa eina töflu yfir öll gögnin og úr því búið til stöplarit.

Einnig vildum við sjá hversu oft þau voru að lenda í þessum svokölluðu “top 3” sætum. Notað var talninguna yfir fjölda keppna hjá þeim og síðan aukalega talið hversu oft þau lentu í Rank númer eitt, tvö eða þrjú. Þeirri talningu var síðan deilt með fjölda keppna og út úr því kom prósentan yfir þau skipti sem þau Bergdís og Eyjólfur lentu í efstu þremur sætunum. Því var síðan sett inn í stöplarit og merkt alla viðeigandi þætti.

Það síðasta sem við skoðuðum var meðaltíminn hjá þeim per kílómeter eða svokallað “pace”. Þar sem þau bæði tvö eru aðallega í hjólreiðum þá voru allar hlaupakeppnir teknar út og aðeins hjólreiða keppnir notaðar. Því næst þurfti að draga úr nafna dálknum yfir heitin á keppnunum fjöldann af kílómetrum í sér dálk. Tíma uppröðunin var HH:MM:SS og því skipt yfir í aðeins mínútur. Sett var svo upp línurit sem sýndi meðaltíma per kílómeter í hverri keppni.