



Gagnavinnsla

T-316-GAVI.
Hópaverkefni 3

Kennari: Eyjólfur Ingi Ásgeirsson eyjo@ru.is

Ágúst Aðalsteinsson agust11@ru.is
Björg Marium Adamsdóttir bjorga13@ru.is
Björgvin Már Guðjónsson bjorgving14@ru.is
Guðmundur Hákon Hermannsson gudmundurh14@ru.is
Valgeir Hrafn Snorrason valgeir14@ru.is

Inngangur

Í þessu verkefni er verið að bera saman innflutningsgögn á ýmsum vöruflokkum sem eru fluttir inn til Íslands. Með þessu er t.d. hægt að ákvarða hvort einhver vöruflokkur sé á undanhaldi gagnvart öðrum vöruflokkum, þ.e. hvort hann gæti haft einhver áhrif á innflutning hans. Einnig er möguleiki í þessu forriti að skoða innflutningsmagn á einum flokk.

Aðferð

Gögnin voru sótt á vefsíðu Hagstofu Íslands og unnum með gögn sem heita: [Innflutningur eftir vöruflokkum, SITC, 4,1999 - 2014.](#)

readCSV python script

Gögnin eru lesin í gegnum pandas read_csv inn í pandas array.

Þar eru dálkunum breytt og vöruflokkaheitin og númerin tekin út til að setja inn í SQL.

Stuðst er svo við tvöfalda forlykkju þar sem öll gögnin eru lesin inn, sleppt var að hafa "Alls" röðina, en notað verður SQL select skipanir til að sækja þau gögn.

Forlykkjan byrjar á að inserta inn í categories töfluna og svo fer í seinni forlykkjuna þar sem gögnin voru sótt með data[1:].iterrows() og látin inn í units töfluna.

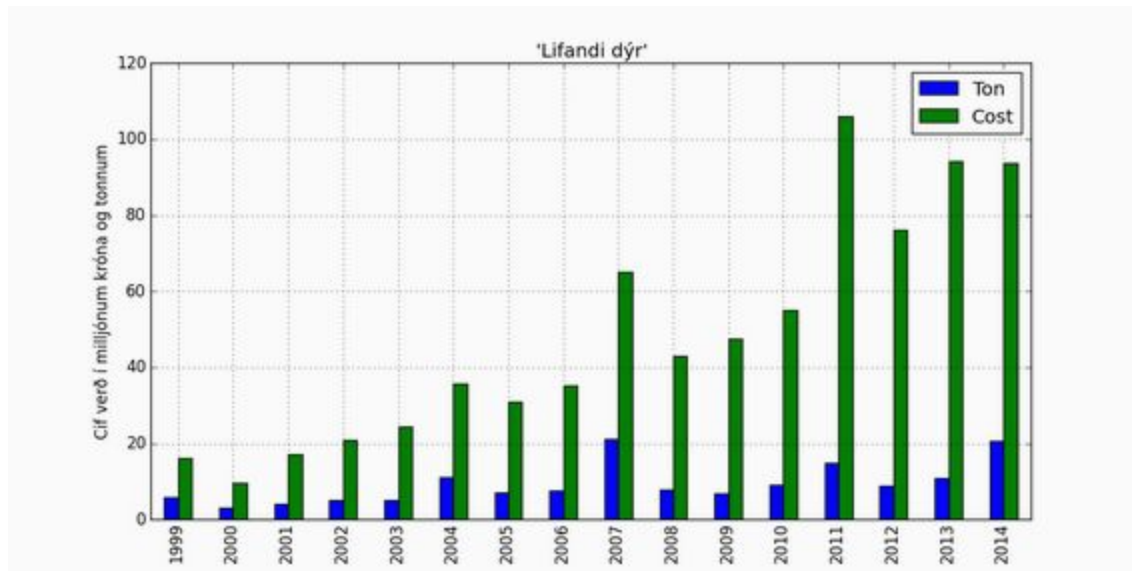
gui python script

gui.py er consoleið þar sem fjórir valmöguleikar birtast:

- 1: Skoða einn vöruflokk
- 2: Bera saman 2 vöruflokka
- 3: Sjá lista yfir vöruflokka
- q: Til að hætta

Valmöguleiki 1: Skoða einn vöruflokk

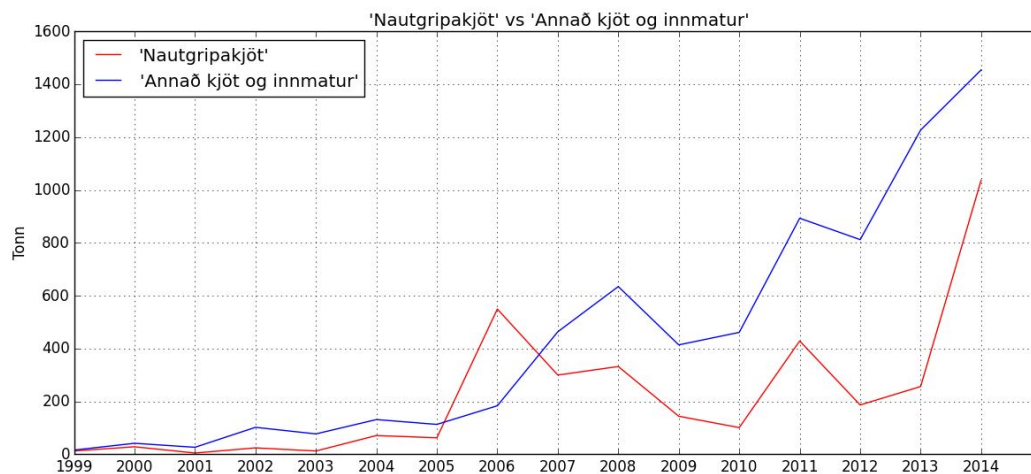
Þegar notandi velur valmöguleikann "Skoða einn vöruflokk" og velur t.d. vöruflokkur 1. sem heitir "Lifandi dýr". Þá prentar forritið töfluna í consoleið með tölunum sem sótt voru frá SQL gagnagrunninum ásamt "Meðaltal þyngdar" og "staðalfrávik" og birtir forritið súlurit sem sýnir á litríkan hátt með hjálp matplotlib.pyplot innflutning eftir þyngd og kostnaði öll árin í gögnunum sem voru sótt frá Hagstofunni. Sjá mynd "Lifandi dýr".



Valmöguleiki 2: Bera saman 2 vöruflokka

Hér byrjar notandi að slá inn fyrsta vöruflokkinn, t.d. 11. sem heitir “Nautgripakjöt” og svo seinni vöruflokkinn í þessu dæmi 12. “Annað kjöt og innmatur”. Svo spyr forritið notendann að slá inn annað hvort 1 eða 2 til að bera saman Tons eða Cost. Forritið birtir svo upp í töflu gögnin sem sótt voru með aðstoð pandas bókasafninu.

Að lokum birtir forritið línurit sem sýnir mismun vöruflokkana eftir árum á lýsandi og áhugaverðan hátt. Sjá mynd “‘Nautgripakjöt’ vs ‘Annað kjöt og innmatur’”

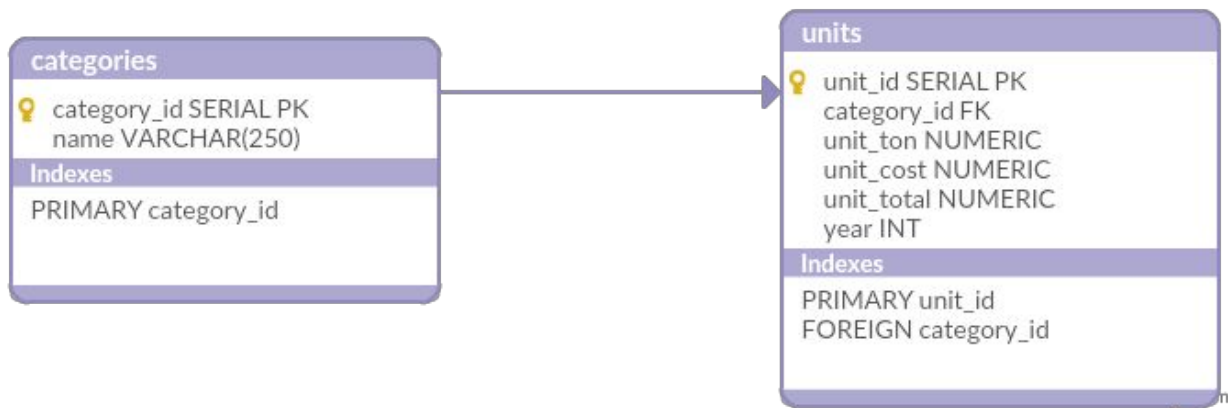


Valmöguleiki 3: Sjá lista yfir vöruflokka

Birtir notanda lista yfir alla vöruflokkana ásamt nöfnunum á þeim.

Gagnaskema

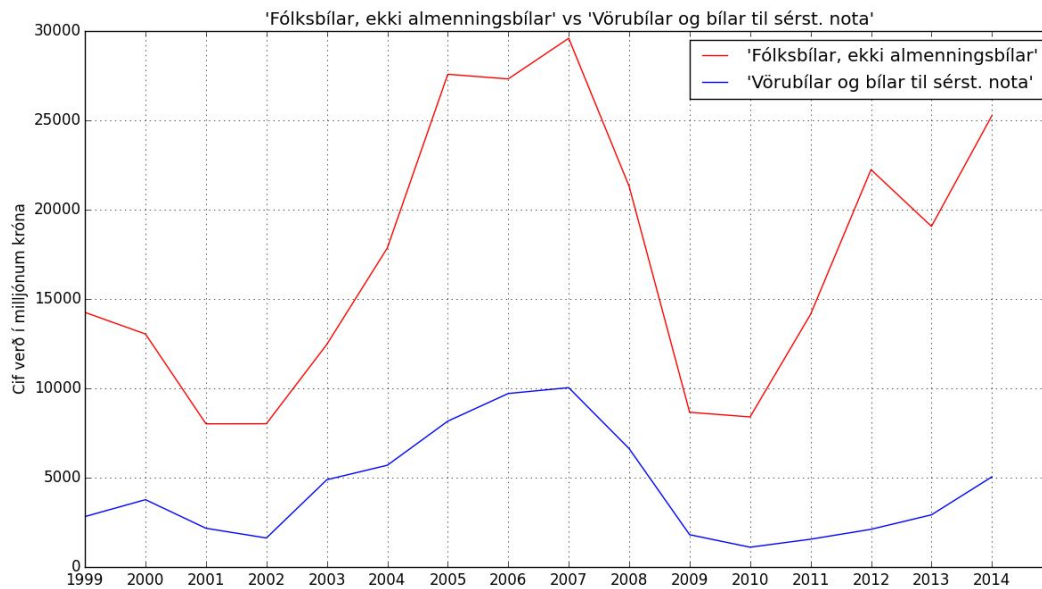
Gagnagrunnurinn geymir tvær töflur categories og units. Taflan categories geymir vöruflokk (name) og auðkenni (category_id) sem er primary lykillinn. Taflan units geymir hversu mörg tonn af ákveðnum vöruflokki voru innflutt eftir árum, hver er kostnaðurinn á vörunum (Cif verð) eftir árum og hlutfall af heild eftir árum. Primary lykillinn í units töflunni er unit_id og svo hefur taflan category_id sem tengir töflurnar saman.



Samantekt

Þar sem þetta verkefni er ekki að leysa/kanna nein vandmál er ekki hægt að setja fram niðurstöður. Best er að sýna dæmi hvernig hægt væri að nota forritið.

Ágætis dæmi sem hægt væri að taka er innflutningur á vöruflokki 781 sem eru fólksbílar og 782 sem eru vörubílar.



Einnig er hægt að sjá innflutning á einum flokki og notum við flokk 054 sem ber heitið Grænmeti nýtt, kælt eða fryst.

