



# Hópaverkefni 1

## Gagnavinnsla

Davíð Halldórsson  
Fannar Örn Arnarsson  
Jóhannes Hilmarsson  
Ómar Sindri Jóhannsson

Dagsetning skýrslugerðar: 4. desember 2015  
Kennari: Eyjólfur Ingi Ásgeirsson

# 1 Rannsóknarspurning

Er eitthvað samband milli afbrotatíðni á Íslandi og auknum mannfjölda? Getum við sýnd fram á það með gögnum frá hagstofunni, með auknum mannfjölda á Íslandi er augljós aukning eða jafnvel fækkun á afbrotum. Er einhver möguleiki á því að afbrotum hafi hlutfallslega ekkert breyst með tímanum?

## 2 Tilgáta

Tilgátann er sú að gögnin sem við erum með segi okkur að afbrot hafa aukist töluvert með tímanum. Tilgátan er sú að við sýnum fram á það að með auknum mannfjölda verða afbrot mun fleiri.



### 3 Forsendur

Við gefum okkur það að eina sem hefur áhrif á afbrot sé mannfjöldi. Kóðinn okkar reiknar ekki með öðrum ástæðum fyrir aukinni eða minnkandi afbrotatíðni sem fræðilega séð geta verið margar.

### 4 Aðferð

Það sem kóðinn gerir er að byrja á því að setja inn python pakkann „Pandas“. Þegar það er búið er hann hannaður til að geta lesið inn nánast hvaða tvær csv skrá frá hagstofunni og bera þær saman. Skilyrðin sem kóðinn setur er að csv skrárnar verða innhalda dálk sem heitir ‘Ár’. Kóðinn gerir dálkinn ‘Ár’ í báðum skjölunum að index til að geta sameinað skjölin. Annað sem kóðinn setur hömlur við er að ef vel er um tegund flokk má aðeins velja einn flokk, en ef velja skal ár mælum við með að taka sem flest ár til að fá sem mesta nákvæmni. Þar á eftir tekur hann þessi tvö skjöl finnur sameiginlegt tímabil sem csv skránar bjóða uppá og setur í eitt „dataframe“.

Þegar þessi uppseting á „dataframe“ er lokið tekur python pakkinn „Numpy“ við. Numpy tekur gögnin og finnur staðalfrávik, meðaltal, miðigildi, hæðsta og lægsta gildið. Þegar þessar reikningar eru búnir þá tekur kóðinn bæði sameiginlega „dataframe-ið“ og öll tölfræði gögn sem „Numpy“ hefur fundið og prentar í tvo csv skrár til frekari greiningar. Í þessu verkefni tókum við tvö csv skjöl frá hagstofunni annars vegar eitt stærsta skjal sem hagstofan býður uppá mannfjöldatalningu frá 1703. Hins vegar skjal sem innheldur tíðni afbrota á Íslandi frá 1999.

```
import csv as csv
import func
```

```
Nameoffile1='mannfjoldi.csv'
Nameoffile2='mannfjoldi.csv'
```

### 5 Niðurstöður