Statistik och dataanalys I, 15 hp

Inlämningsuppgift 2

Namn 1

Namn 2

Namn 3

2/3/23

Innehåll

|  |
| --- |
| Varning |
| Den här inlämningsuppgiften förutsätter att följande paket finns installerade:   * mosaic   Paket kan installeras via kommandot install.packages('packagename'), där 'packagename' är namnet på paketet, t.ex 'mosaic'. |

## Introduktion

I den första inlämningsuppgiften ska ni självständigt i grupper om tre analysera ett dataset i programmeringsspråket R. Till skillnad från datorlaborationerna finns det minimalt med kodexempel. Datorlaborationerna går igenom de flesta momenten som behandlas i inlämningsuppgiften, så se till att göra klart dessa innan.

|  |
| --- |
| Instruktioner |
| I denna inlämningsuppgift ska ni analysera ett datamaterial … Följande variabler finns i datasetet boston\_census\_data.Rdata ([ladda ner](https://github.com/StatisticsSU/SDA1/blob/main/assignments/assignment1/Boston_census_data.RData?raw=true)) …   * town: Stadsdel. * longitude: Longitud koordinat.. * lower\_stat\_pct: Procentandel underklass definerad som en av två: (i) andel vuxna utan gymnasieutbildning eller (ii) andel män som genomför okvalificerat arbete.   ROADMAP  Inlämningsuppgiften ska lämnas in i form av ett html dokument genererat av Quarto. **Kontrollera noga att du inte har några felmeddelande och att dokumentet kompileras utan problem**. Använd tydliga figurer och namnge axlarna med tydliga variabelnamn. Glöm inte att skriva era namn ovanför istället för Namn 1, Namn 2 och Namn 3. |

## 0. Ladda in data

#### 💪 Uppgift 0.1

Ladda in dataseten Boston\_census\_data.Rdata och Boston\_districts\_to\_predict.Rdata (länkar för att ladda ner data finns i Instruktioner avsnittet ovan).

|  |
| --- |
| **Uppgift 0.1 - Svar**  # Write your code here load(file = url("https://github.com/StatisticsSU/SDA1/blob/main/assignments/assignment1/Boston\_census\_data.RData?raw=true")) |

## 1. Kriminalitet i Boston

I detta avsnitt ska ni analysera kriminaliteten i Boston med hjälp av variabeln crime\_rate.

#### 💪 Uppgift 1.1

Vad kan man generellt säga om kriminaliteten i censusdistrikten? Använd lämpliga figurer samt fördelningsmått som underlag.

|  |
| --- |
| **Uppgift 1.1 - Svar**  Skriv svaret här. Vid behov skrivs matematiska symboler inom dollartecken, till exempel . Koden skrivs i R-rutan nedanför.  # Write your code here |

#### 💪 Uppgift 1.2

Varierar brottsligheten i Boston beroende på den kategoriska variabeln town? Det finns 88 olika utfall av town (dvs 88 olika stadsdelar). Välj ut Boston East Boston, Boston Downtown,Cambridge, samt två valfria stadsdelar för att besvara frågan. Frågan besvaras med hjälp av lämpligt valda figurer och statistiska mått.

|  |
| --- |
| Tips |
| Skapa en ny data frame som filtrerar Boston\_census\_data (till exempel genom filter() funktionen) utefter de stadsdelarna ni är intresserade utav innan ni påbörjar analysen. |

|  |
| --- |
| **Uppgift 1.2 - Svar**  Skriv svaret här.  # Write your code here |

#### 💪 Uppgift 1.3

Vilka tre variabler i datasetet Boston\_census\_data korrelerar mest med brottslighet? Beskriv det parvisa sambandet mellan brottslighet och vardera av dessa tre variabler.

|  |
| --- |
| Tips |
| Kom ihåg att korrelation är ett beroendemått för *numeriska variabler*. |

|  |
| --- |
| **Uppgift 1.3 - Svar**  Skriv svaret här.  # Write your code here |

## 2. Fastighetsskatt i Boston

I detta avsnitt ska ni analysera fastighetsskatten i Boston med hjälp av variabeln tax\_rate.

#### 💪 Uppgift 2.1

Vad kan man generellt säga om fastighetsskatten i censusdistrikten? Använd lämpliga figurer samt fördelningsmått som underlag.

|  |
| --- |
| **Uppgift 2.1 - Svar**  Skriv svaret här.  # Write your code here |