

ggplot2 - Oefening

Maak je vertrouwd met de `airquality` data in R

1. Bekijk de help file van `airquality`.
2. Welke variabelen zitten in de dataset?
 - a. Van welk datatype zijn deze variabelen?
3. Vraag een `summary()` van de gegevens.
 - a. Voor welke maanden beschikken we over gegevens?
 - b. In welke eenheid is de temperatuur gemeten?
 - c. Zijn er ontbrekende waarden voor bepaalde variabelen?

Verkennde plotjes

Wees aandachtig voor warnings en foutmeldingen. Probeer deze te begrijpen alvorens verder te gaan met de oefening.

1. Bekijk de verdeling van de temperatuur.
 - a. Maak hiervoor een histogram.
 - b. Verander `binwidth = 5`.
2. Bekijk de verdeling over de maanden.
 - a. Maak hiervoor een barplot.
3. Verschilt de gemeten temperatuur per maand?
 - a. Maak hiervoor een boxplot.
 - b. Splits deze op volgens maand.
 - c. Krijg je de figuur die je in gedachten had? Waarom (niet)? Lees zeker de warning.
 - d. Hoe kan je dit oplossen? (Denk aan het datatype)
4. Is er een verband tussen temperatuur en ozon?
 - a. Maak hiervoor een scatterplot.
 - b. Voeg een (al dan niet lineaire) smoother toe.
 - c. Wat betekent de warning? Moeten we hiervoor iets aanpassen?
 - d. Verander de titel in “Vernand tussen temperatuur en ozon”.
 - e. Kies duidelijke namen (met eenheid) voor de X- en Y-assen.
5. Is het verband tussen temperatuur en ozon afhankelijk van de straling van de zon?
 - a. Voeg aan de vorige scatterplot een kleur toe voor zonnestraling.
 - b. Kunnen we nu ook afzonderlijke smoothers voor de zonnestraling bepalen? Waarom (niet)?
6. Is het verband tussen temperatuur en ozon afhankelijk van de maand?
 - a. Maak hiervoor per maand een scatterplot met een rode lineaire smoother.
 - b. Bewaar deze figuur in de map “Figuren/” onder de naam “TempOzon_PerMaand.png”.