

Övningsuppgift klustering

Gör en funktion, `data_generator`, med följande inargument:

- `n_vals` – antal värden som returneras
 - `x_min` – minsta värde för `x`
 - `x_max` – maximala värde för `x`
 - `y_min` – minsta värde för `y`
 - `y_max` – maximala värde för `y` (hör och häpna)
 - `r_min` – minimala radie ($r_{\min} > 0$)
 - `r_max` – maximala radie ($r_{\max} < \sqrt{x^2 + y^2}$)
 - `tol` – toleransparameter
-
- `Data_generator` ska generera värden för `x`, `y` från en uniform distribution från intervallet givet av `x`- och `y`-parametrarna.
 - Returnera en dataframe (as_tibble) som har filtrerat på två olika sätt:
 - Ta bort alla datapunkter som ligger i ett cirkelintervall [`r_min`, `r_max`]. Dvs behålla datapunkter $\sqrt{x^2 + y^2} < r_{\min}$ och $\sqrt{x^2 + y^2} > r_{\max}$. Tips använda funktionen `filter`.
 - Filtrera dataframen med ett kryss. Dvs filtrera bort `x + y` i intervallet `[-tol, tol]` och filtrera bort `x - y` i intervallet `[-tol, tol]`.

Kalla på funktionen och plotta dina resultat med `ggplot` där du anger vad de olika parametervärden ska vara.

Testa använda DBScan för att klustra på datasetet du nu har skapat och plotta resultatet ihop med datasetet.