

Angst & depression

Folkesygdomme

Angst er en normal reaktion på situationer, der er truende eller farlige, men for nogle er angsten vedvarende og sygelig. Her taler man om "generaliseret angst". Symptomer omfatter overdreven bekymring, frygt og fysisk ubehag som muskelspænding og hjertebanken.

Depression er en psykisk lidelse, som påvirker humør, tanker og kropsfunktioner. Symptomer omfatter vedvarende tristhed, manglende interesse eller glæde samt søvnforstyrrelser.

Behandlingen for begge sygdomme indebærer ofte samtaleterapi og medicin. Med rettidig behandling er prognosen for angst og depression sædvanligvis god. De fleste oplever en forøget livskvalitet efter noget tid og nogle kommer sig fuldstændigt.

PHQ (Patient Health Questionnaire) og **GAD** (Generalized Anxiety Disorder) er screeningsværktøjer, der bruges til at vurdere alvorligheden af depression og angst. En høj scorer indikerer alvorlige symptomer, og disse scorer bruges af sundhedspersonale til at diagnosticere og træffe behandlingsbeslutninger.

PHQ og GAD	Sværhedsgrad
0-4	minimal
5-9	mild
10-14	moderat
15-19	moderat-svær
20-27	svær

Sammenhæng mellem angst og depression

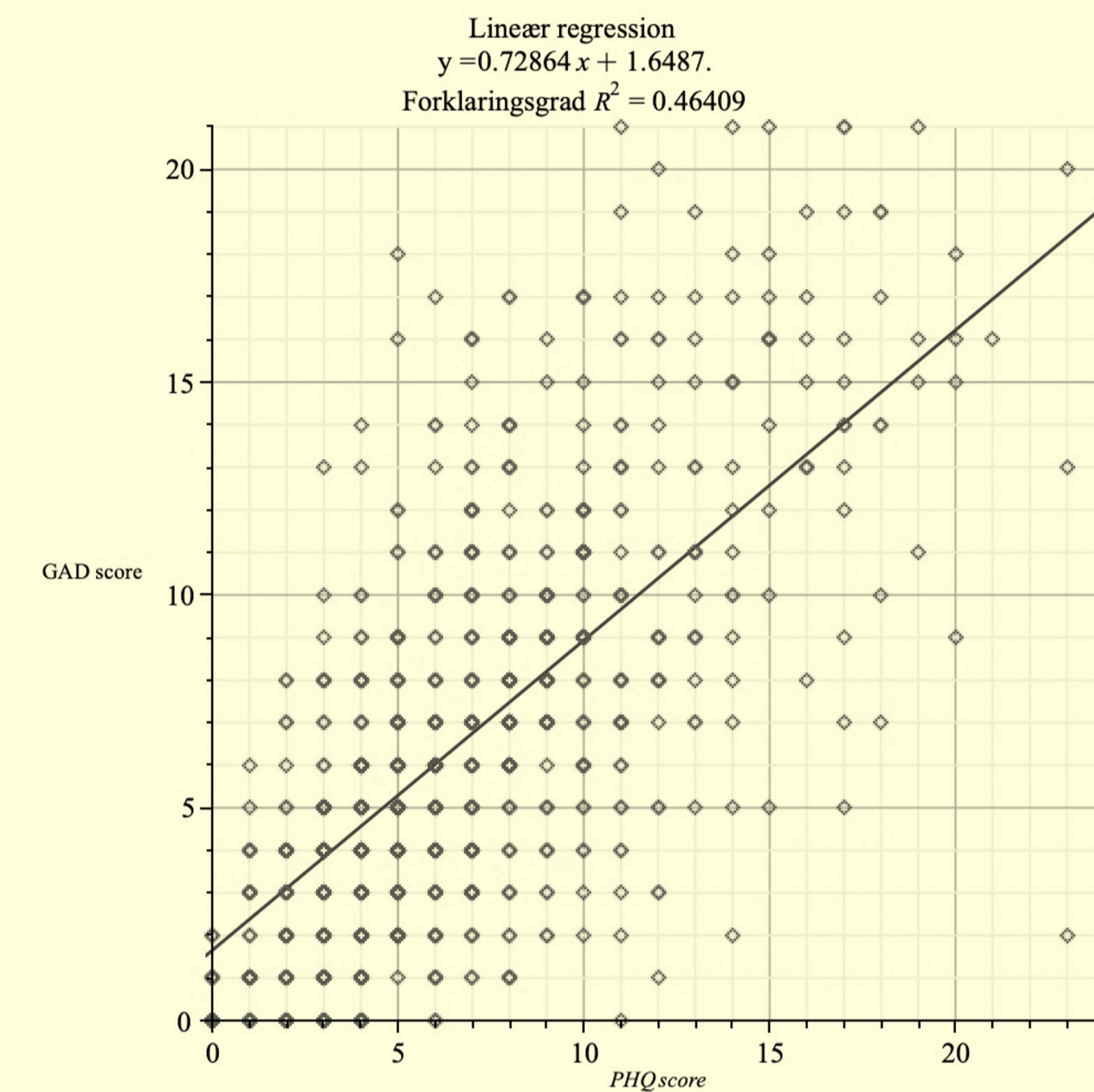
Angst og depression er ofte komorbide, hvilket betyder, at de forekommer samtidigt og urelateret, og dette kan komplicere patientens tilstand. Hvis en patient både har angst og depression, er dette vigtigt at tage i betragtning i behandlingen for at optimere resultater og forbedre patientens livskvalitet.



Her ses et punktdiagram med de to kontinuerte variable PHQ og GAD score. Tendenslinjen har forskriften:

$$y = 0.73x + 1.65$$

Hældningskoefficienten 0.73 beskriver, hvor meget GAD scoren stiger, når PHQ scoren stiger med én. Konstantleddet 1.65 angiver GAD scoren, når PHQ scoren er lig med nul.



Kausalitet eller korrelation?

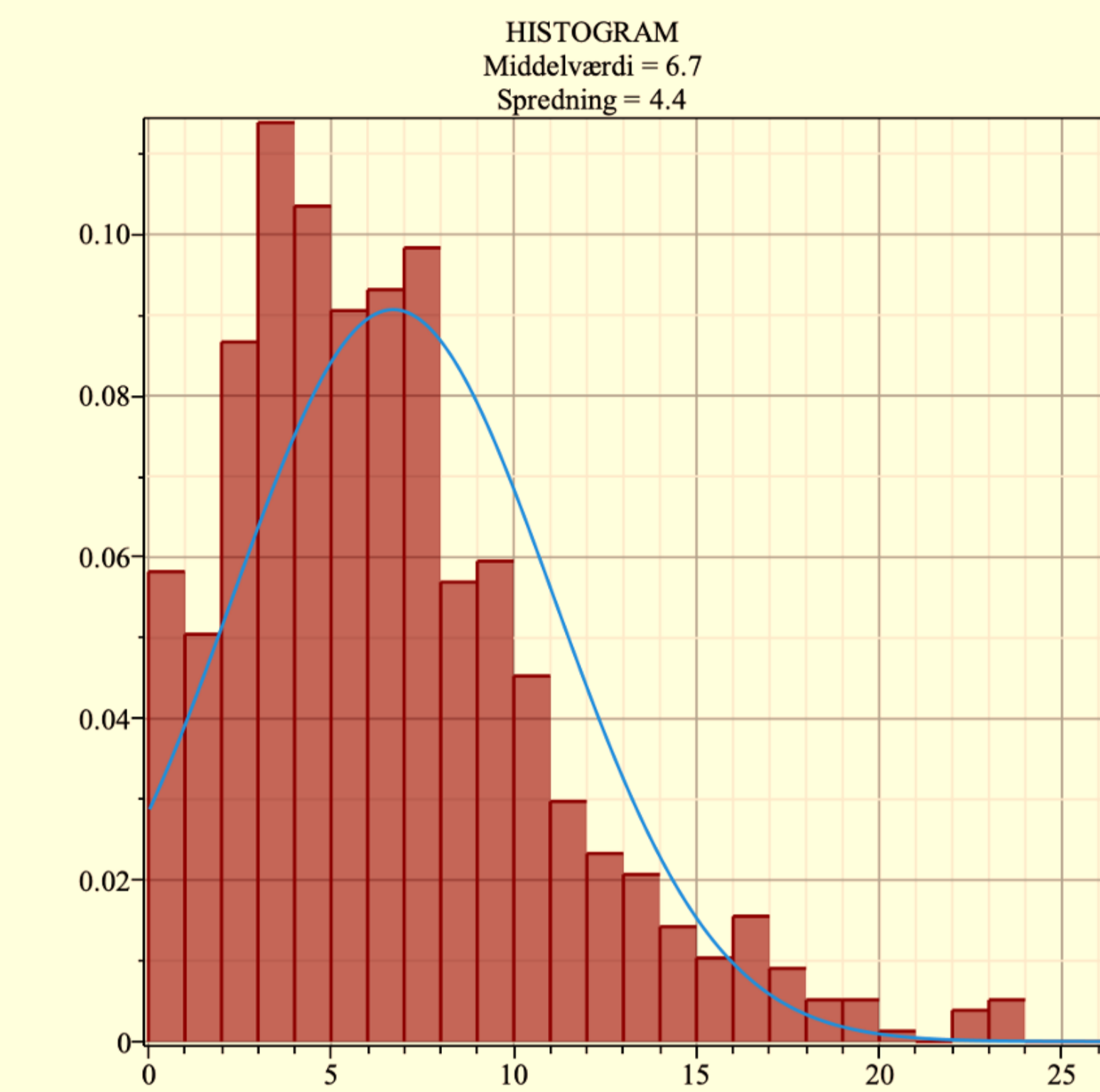
Hvis man betragter andelen, der tager antidepressiv og har selvmordstanker, kunne man konkludere at antidepressiv forårsager selvmordstanker. Dette er dog en logisk fejlslutning. Sammenhængen skyldes formentlig korrelation og ikke kausalitet.

Kausalitet angiver en direkte årsagssammenhæng, mens korrelation kan skyldes en tilfældighed eller en bagvedliggende fællesårsag fx at individer med depression oftere er selvmordstruede og tager antidepressiv.

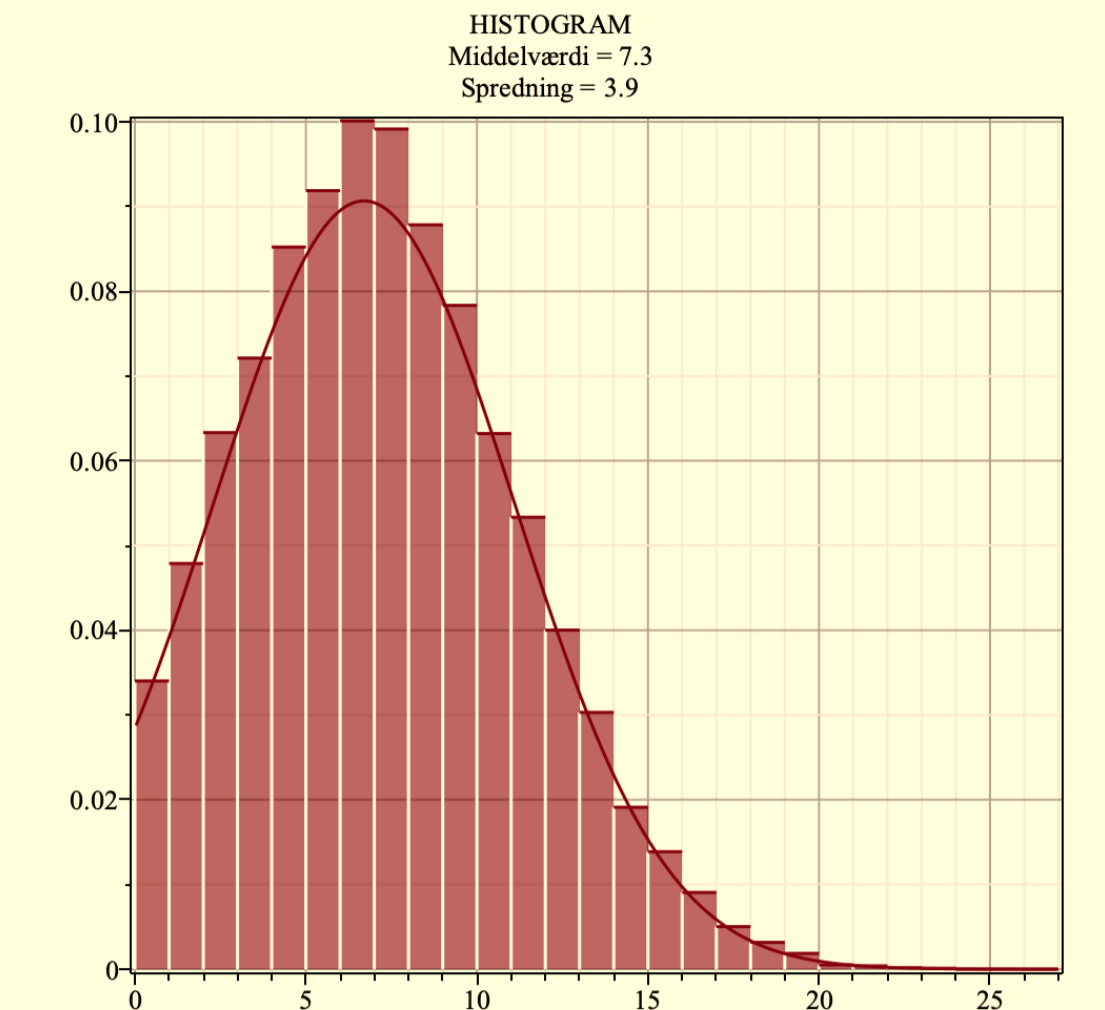
Da angst og depression ofte er komorbide, er den lineære sammenhæng mellem dem formentlig korrelation.

Er PHQ scoren normalfordelt?

Plotter man et histogram for PHQ scoren sammen med en gaussisk kurve, ser man at det tilnærmelsesvist ligner en normalfordeling. Der er dog væsentlige afvigelser i det ekstraordinære område. Ved optælling har 1.54% af deltagerne svær depression. Når vi udregner denne andel for en normalfordeling med samme middelværdi og spredning, får vi 0.26% i denne kategori. Dette er en væsentlig forskel og kan tyde på at vores data ikke er normalfordelt!



Simulerer man et normalfordelt datasæt med 19575 værdier med middelværdien 6.7 og spredningen 4.4, og plotter dette som et histogram bliver den nye middelværdi 7.3 og spredningen 3.9. Dette skyldes, at det nye datasæt er normalfordelt. Sandsynligheden for at have svær depression i dette datasæt er 0.135%, hvilket ligger væsentligt under den observerede andel.



Er vores data repræsentativ?

Vi vil gerne kunne benytte vores data til at sige noget om den generelle danske befolkning. Da vores data stammer fra studerende på University College of Lahore i Pakistan er det vanskeligt at sige noget om forekomsten eller fordelingen af angst og depression hos den danske befolkning. Hvis vi antager, at den danske befolkning er fordelt ligesom de forskellige datasæt får vi følgende værdier i kategorien "svær depression":

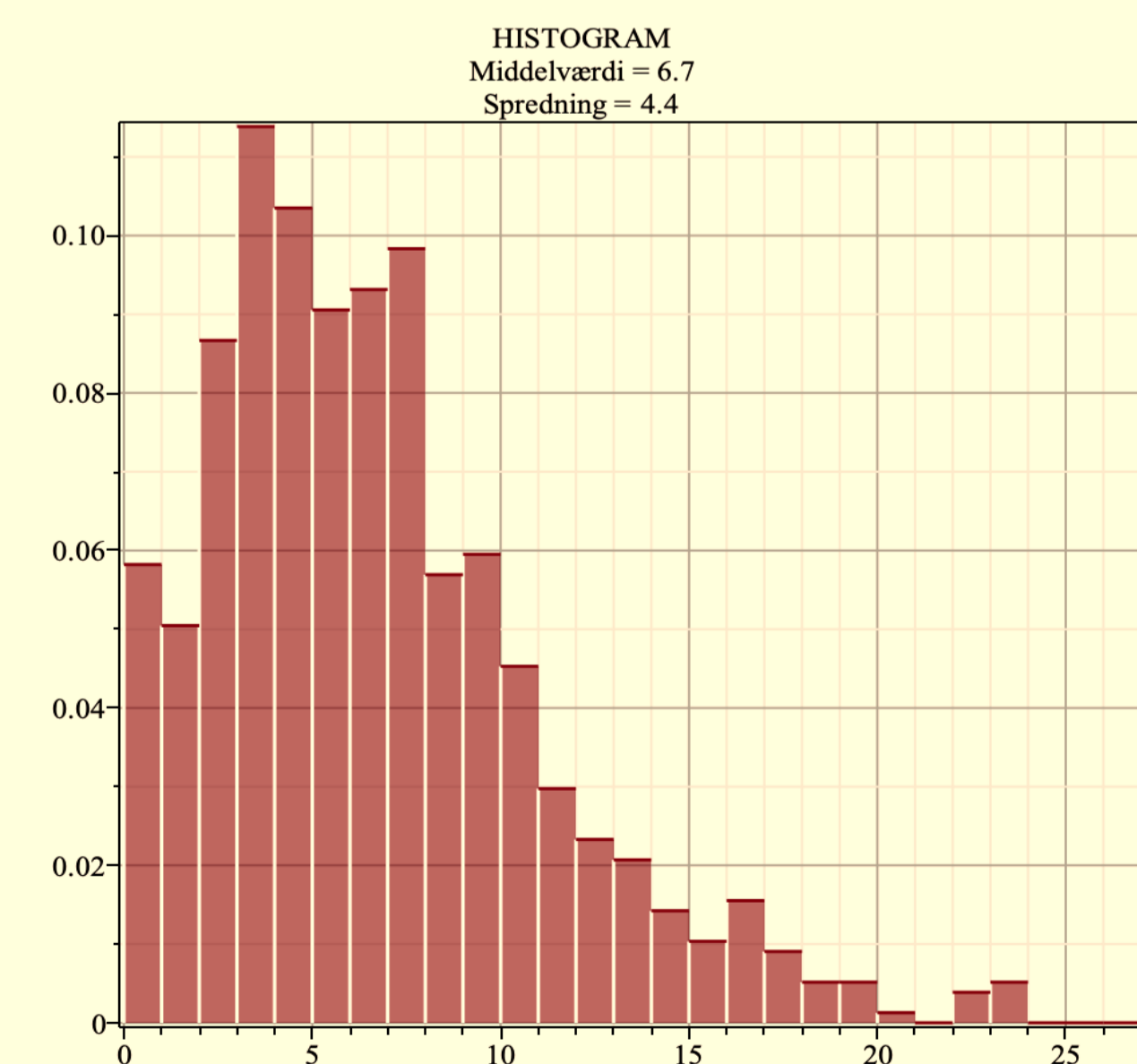
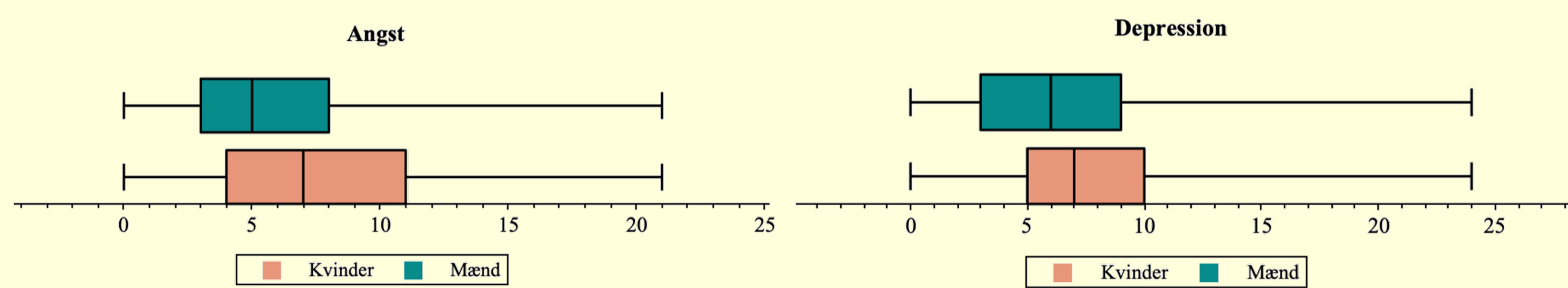
Det oprindelige datasæt: ~ 90200
Normalfordeling med $\mu = 6.7$ og $\sigma = 4.4$: ~ 15200
Normalfordeling med $\mu = 7.3$ og $\sigma = 3.9$: ~ 7900

Vi kender ikke antallet af danskere med svær depression, men vi ved, at 91.000 danskere årligt er i kontakt med psykiatrien i forbindelse med depression. Da behandling i psykiatrien ofte kun bevilges dem, som er hårdest ramt, er det rimeligt at antage, at ca. 91.000 danskere har svær depression. Dette stemmer overens med andelen af studerende med svær depression i vores datasæt.

For at kunne konkludere noget mere sikkert om forekomsten af depression i den danske befolkning ville det dog kræve et datasæt fra Danmark med patienter med varierende aldre, baggrund, geografi mm.

Data

Vores datasæt indeholder information om 783 studerende på University College of Lahore i Pakistan. Dataet omfatter en række diskrete og kontinuerte variable, herunder bl.a. køn, BMI, Epworth (søvnproblemer) PHQ og GAD. Vi vil prøve at undersøge skjulte sammenhænge mellem disse variable.



De 2 boksplots viser kvartilsættene for PHQ og GAD score for den diskrete variabel "køn". For begge sygdommene gælder det, at kvindernes kvartiler er højere. For PHQ har kvinderne en median på 7, mens mændenes er 6, og for GAD har kvinderne en median på 7 og mændenes er 5. Boksplottene viser således, at kvinder generelt har en højere PHQ og GAD score.

Histogrammet viser fordelingen for PHQ scoren. Middelværdien er 6.7 og spredningen er 4.4. Senere undersøges, om dataet er normalfordelt.

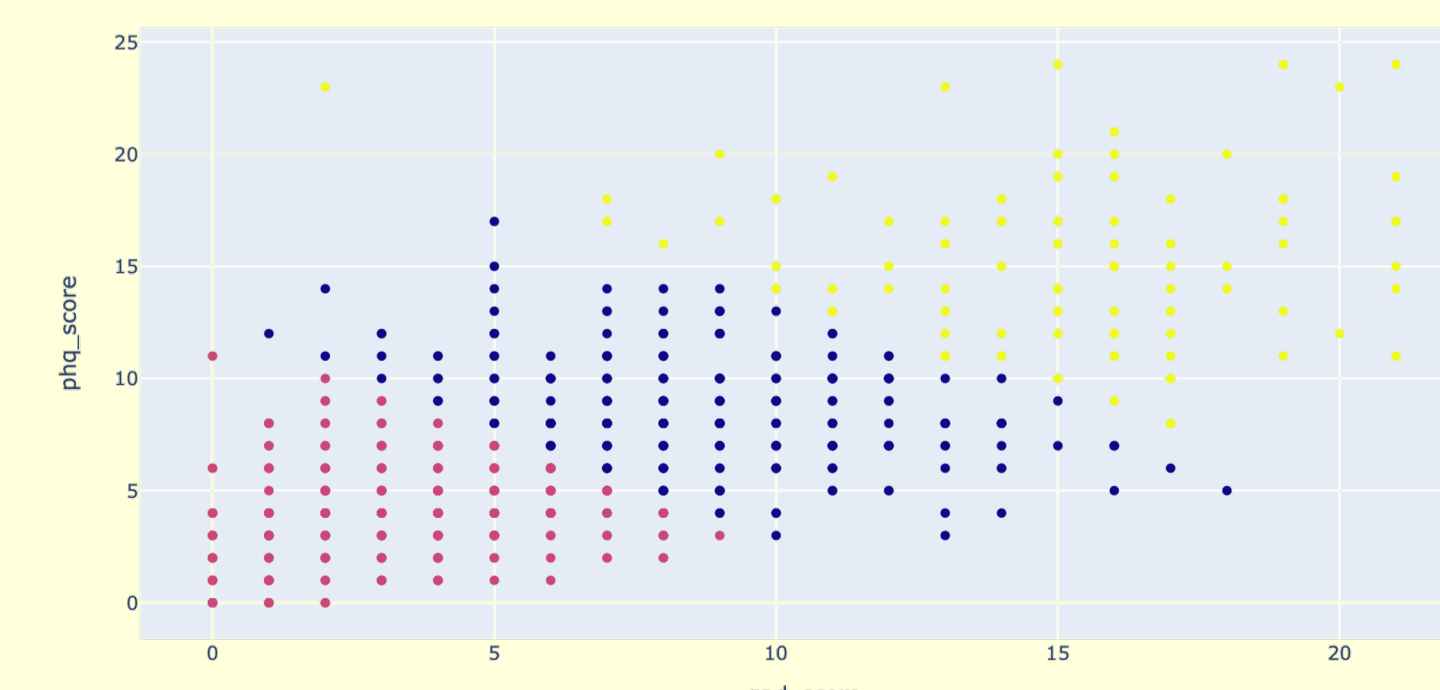
K-means clustering

Princippet

K-means clustering er en metode til at gruppere data i såkaldte clusters. Et **cluster** er en gruppe af datapunkter med lignende egenskaber. Hvert cluster har en **centroid** som er midtpunktet i clusteret. Et datapunkt tildeles det cluster, som det har den mindste euklidiske afstand til. Derefter definerer man en ny centroid, som er gennemsnittet for det pågældende cluster, og den euklidiske afstand til hvert datapunkt beregnes igen, hvorefter hvert datapunkt tildeles et nyt cluster. Denne proces gentages indtil centroiderne ikke ændrer sig mere.

'Unsupervised learning' er en machine learning teknik, som bruges til at finde mønstre i et datasæt. På den måde kan man undersøge interessante tendenser.

Her ses fx et punktdiagram for PHQ og GAD score, som er inddelt i 3 clusters.



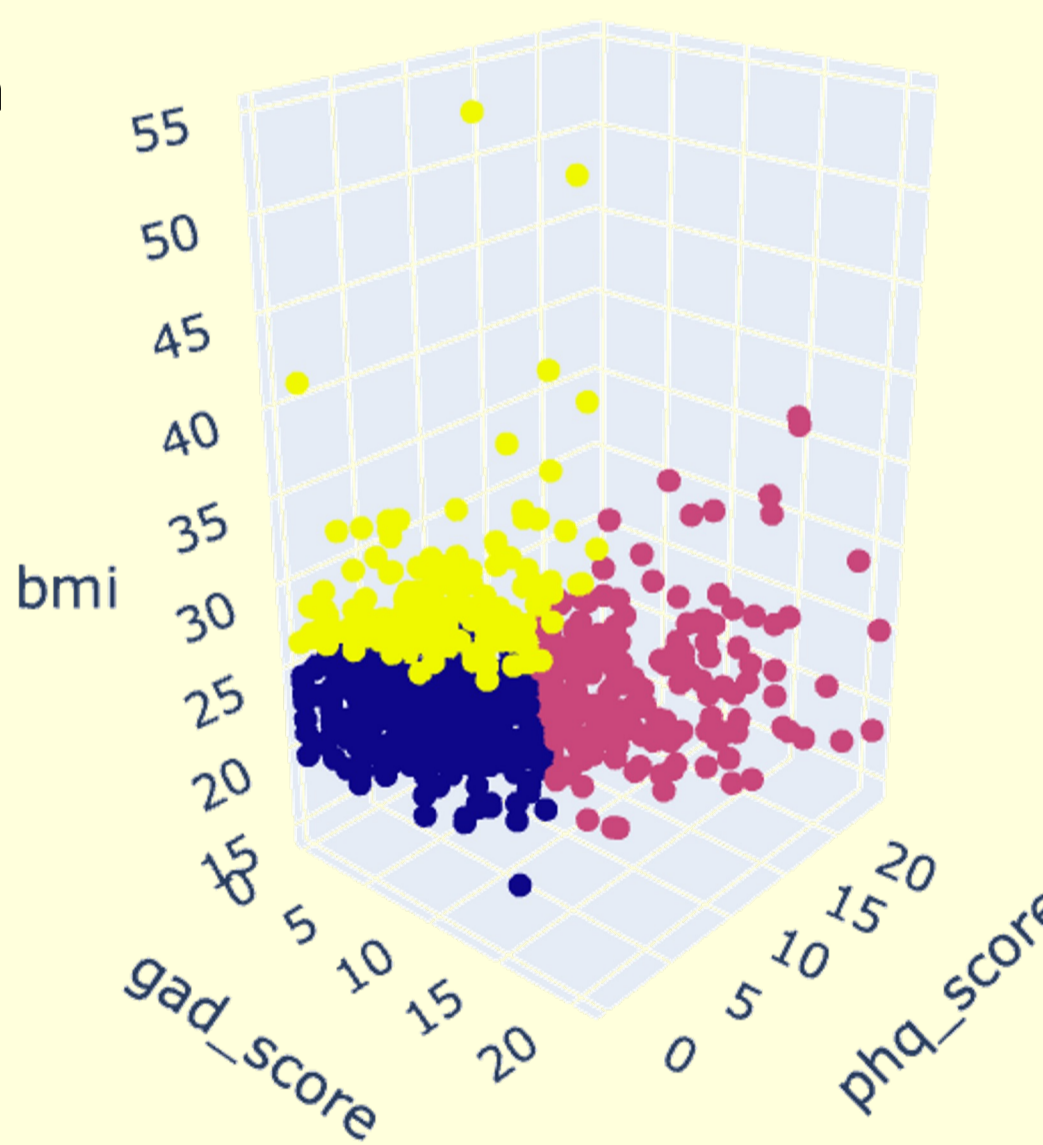
Vha. af 'Unsupervised Learning' har vi undersøgt sammenhængen mellem forskellige variable og bestemt 3 clusters.

BMI

Det er interessant at undersøge sammenhængen mellem BMI (Body Mass Index), som angiver sammenhængen mellem en persons vægt og højde og forekomsten af angst og depression, da den fysiske sundhed har en stor indflydelse på ens mentale sundhed, og BMI kan være en *indikation* på ens mentale sundhed.

Her har vi undersøgt sammenhængen mellem PHQ, GAD og BMI og bestemt 3 clusters.

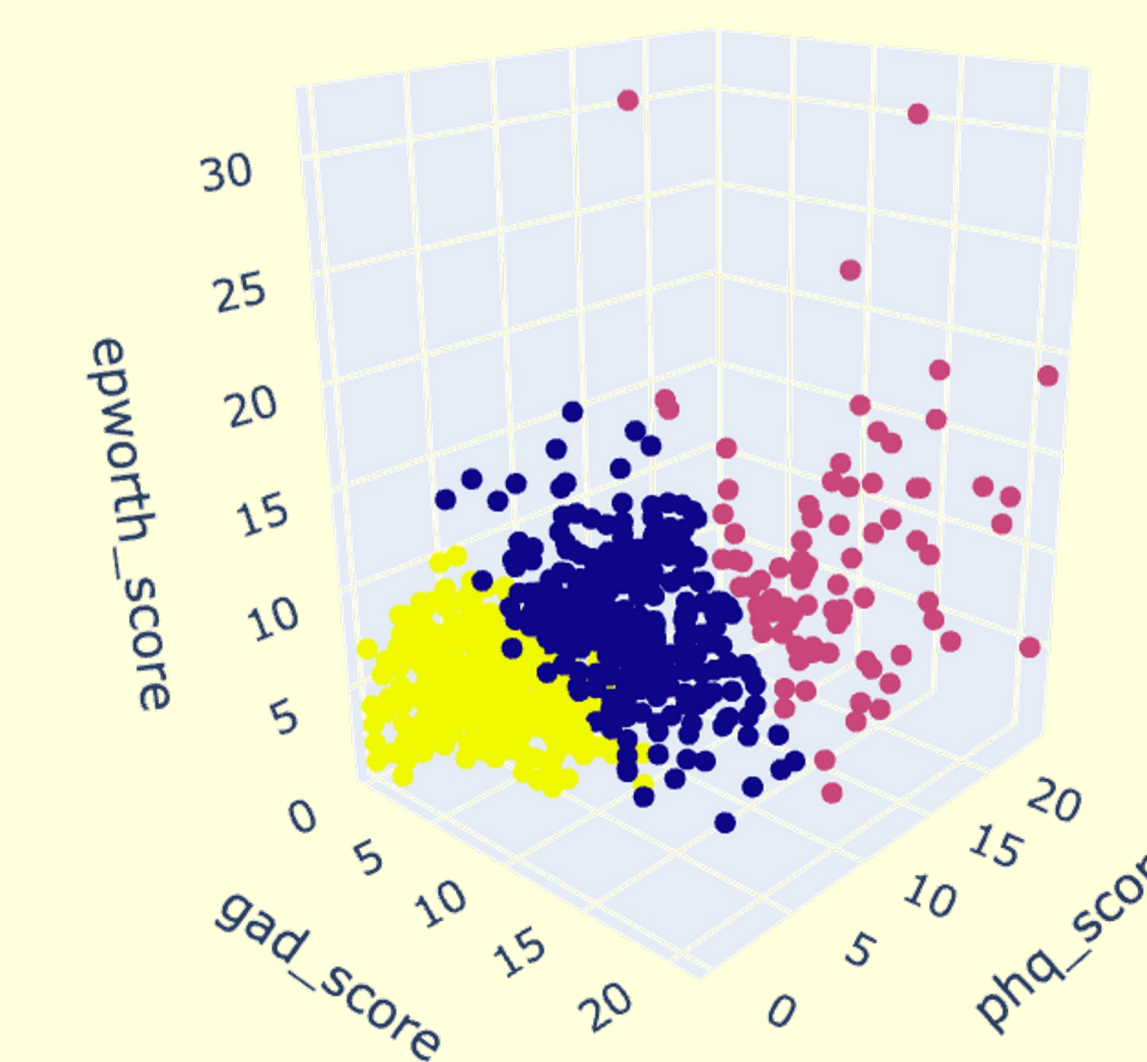
Det gule og det blå cluster ligger ved en lav PHQ og GAD score og de indikerer 2 grupper af studerende, der har en hhv. lav og høj BMI. Det røde cluster ligger ved en høj PHQ og GAD score og har en meget varierende BMI. Dvs. de studerende med en høj PHQ og GAD score har ikke nogle fællestræk ift. BMI. Man kan derfor ikke slutte en sammenhæng mellem BMI og risikoen for angst eller depression.



Søvnproblemer

Søvnproblemer er et hyppigt symptom hos personer med angst og depression. Vi har derfor undersøgt sammenhængen mellem PHQ, GAD og Epworth, som er et udtryk for søvnproblemer.

Ud fra de 3 clusters ligner det, at der er en lineær sammenhæng mellem de 3 variable. Det gule cluster angiver lav PHQ, GAD og Epworth, mens det blå cluster angiver lidt højere scorer. Det røde cluster er karakteriseret ved høj PHQ og GAD score, og her er Epworth scoren meget varieret og indeholder alle de ekstreme tilfælde. Dette tegner umiddelbart en sammenhæng mellem angst/depression og søvnproblemer.



Konklusion & sammenhænge

Udover at have fundet en stærk korrelation mellem angst og depression har vi undersøgt sammenhængene mellem **PHQ**, **GAD** og **køn**, **BMI** og **Epworth**.

Vi kan konkludere, at kvinderne generelt har betydeligt højere PHQ og GAD end mændene. I vores datasæt har vi ikke fundet nogen sammenhæng mellem BMI, PHQ og/eller GAD. Til gengæld kan se en sammenhæng mellem PHQ, GAD og Epworth, hvilket stemmer overens med, at søvnproblemer er et hyppigt symptom hos patienter med angst og depression.

Vores data ser ikke ud til at være normalfordelt, da dette ville indikere, at der var langt færre tilfælde med svær depression, end der i realiteten er. Derudover stammer dataet fra en pakistansk studerende. Derfor er det vanskeligt at sige noget om forekomsten i den generelle danske befolkning ud fra vores datasæt. Her ville vi foreslå at analysere et datasæt med informationer om danskere.

Løsningsforslag

Den bedste løsning til at bekæmpe folkesygdommene, angst og depression, er rettidig og effektiv behandling. Dette omfatter bl.a. samtaleterapi og medicinering. På baggrund af dataet kan man ikke konkludere noget om, *hvilken* behandling der er mest effektiv. Dertil skulle man bruge længerevarende kontrolerede forsøg med forskellige behandlingsformer.

Litteraturliste

Veddel Kessing, L. (2024, 15. Januar). *Depression, en oversigt*. Sundhed.dk. <https://www.sundhed.dk/borger/patienthaandbogen/psyke/sygdomme/depression/depression-en-oversigt/>
Ernst Nielsen, R. (2022, 7. December). *Angst*. Sundhed.dk. <https://www.sundhed.dk/borger/patienthaandbogen/psyke/symptomer/angst/>
Flachs, E.M et. al. (2015). Statens Institut for Folkesundhed, Syddansk Universitet. *Sygdomsbyrden i Danmark - sygdomme*. København: Sundhedsstyrelsen. pp. 173.
Gajawada, S. K. (2019, 13. Maj). *K-Means Clustering with Math - Towards Dat. Science. Medium*. <https://towardsdatascience.com/k-means-clustering-for-beginners-2dc7b2994a4>