ACTU PS5841 Data Science in Finance & Insurance - Autumn 2021 (Y. Wang)
Assignment - 7
Assigned 11/09/21, Due 11/16/21 Tuesday

Problem 1. Poisson Regression
The file data.csv contains accident counts for 5 age groups, along with the corresponding exposure. You decided to use a Poisson regression fit the data:

$$\ln(E(\text{Accidents}_i)) = \ln(\text{personyears}_i) + \beta_0 + \beta_1 \text{Age}_i$$

A. Fit the data using sklearn and report the following:
a1. $[\hat{\beta}_0, \hat{\beta}_1]$
a2. The corresponding log-likelihood value.

B. Fit the data using statsmodels and report the following:
b1. $[\hat{\beta}_0, \hat{\beta}_1]$
b2. The corresponding log-likelihood value
b3. The corresponding pseudo R^2 when compared to the naive model

$$\ln(E(\text{Accidents}_i)) = \ln(\text{personyears}_i) + \beta_0$$