

Team Members : Bao Nguyan, Guillermo Leija, Zain Ul Haq
Group Name :
Name : Seyedeh Marzieh Hosseini – shosseini@uni-potsdam.de
Country: Germany
Specialization : Data science

Business Problem

One of the long-lasting business issues in the world of Pharmaceutical companies is the persistency of drugs which can significantly affect the outcome of medical treatments. One of the important factors that is related to persistency is the adherence of the patient to the prescribed regimens, meaning if the patient is committed to the prescribed regimens or not. In fact, related studies show that around 50%-60% of the patients with different illnesses in US miss doses, take the wrong doses, or drop off treatment in the first year. ([reference 1](#))

Additionally, the illness, either chronic or acute can be related to the adherence and persistency of drugs. To solve the problem, we need to find out factors that are the most correlated to the persistence of drugs and adherence.

Data understanding

The data consists of 3424 rows and 69 columns. There are five group of features, patient information such as age, race, region, gender and Ethnicity. There is another group called provider Attributes. Another group called Clinical factors which describes the clinical situation of the patient. The fourth group of feature is about the risk factors that the patient might undergoes for chronic and acute diseases. Another important feature in this group is the Adherence which describes the adherence of patient to the prescribed regimens.

Project Cycle

Tasks	08/11/2021 Week 0	15/11/2021 Week 1	22/11/2021 Week 2	29/11/2021 Week 3	6/12/2021 Week 4
Week 7					
Week 8					
Week 9					
Week 10					
Week 11					
Week 12					

References :

1- Hayden B. Bosworth et al., "Medication adherence: A call for action," *American Heart Journal*, 2011, Volume 162, Issue 3, pp. 412–24, [ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov/).

Data Intake Report

Name: Health care Dataset

Report date: 5 November 2021

Internship Batch: 0076B8

Version:1.0

Data intake by:Seyedeh Marzieh Hosseini

Data intake reviewer: Bao Nguyen

Data storage location: https://github.com/Khanhbao8695/HealthCar_DS2021

Tabular data details:

Total number of observations	3424
Total number of files	1
Total number of features	69
Base format of the file	xlsx
Size of the data	920KB

GitHub Repository:

Project Link: https://github.com/Khanhbao8695/HealthCar_DS2021