

# Drug Persistence Classification

Created by Pharmalytics

## Week 9: Data Cleansing and Transformation

Noah Igram

[noahigram2023@u.northwestern.edu](mailto:noahigram2023@u.northwestern.edu)

Northwestern University, Evanston IL, USA

Group name: Pharmalytics

ABC Pharma is a pharmaceutical company looking to improve their patient identification process. They hope to improve their reach to customers by targeting more specific groups of people who they believe would be potential drug patients. We hope to automate the process of predicting drug persistence by creating a classification model which uses past patient data to predict whether or not a patient was persistent in taking a drug for its prescribed course.

For data cleansing and transformation, two methods were used. The first method removed any rows in the data set with 'Unknown' or 'Other' as their entry. The second method replaced these entries with the mode of their column. It turned out that the second method yielded better accuracy with our model. The following GitHub link is for a repo which contains the Data Cleansing python notebook as well as all of the other project deliverables.

GitHub Repo: <https://github.com/noahigram/Pharmalytics>