Chair of Decision Sciences and Systems
TUM School of Computation, Information and Technology
Technical University of Munich

TUM

Business Analytics & Machine Learning
Tutorial sheet 6: Data Preparation

Prof. Dr. Martin Bichler
Julius Durmann, Markus Ewert, Yutong Chao, Dr. Mete Ahunbay

Exercise T6.1 *Developing a music recommender system*

Imagine you want to develop a recommender system that suggests new songs to users. To achieve this, your organization collected data on the user's behavior on the platform and whether they favor songs. Your task is to first clean the data sets individually and then merge them into a single data set for the development of the predictive model. Use the merged data sets to answer the questions noted in the provided notebook.

Attribute	Description	Comment
uid	A unique identifier per user	
p	Indicates whether a user has the status *Premium*	nominal: 0 = False, No = False, 1 = True, Yes = True
m1, m2, m3	The minutes listened to music per user for the last three months	

Table 1 Attributes of users.csv

Attribute	Description	Comment
user_id, song_id, artists	A unique identifier per user, song, artist	
num_clicks	The number of times a user clicked on a song	
ml	The minutes a user listened to a song	
g	The genre associated with that song	nominal
f	Indicator whether a user liked a song or not	0 = No Like, 1 = Like
mod	Date at which the user changed their review	

Table 2 Attributes of users_behavior.csv

Attribute	Description	Comment
artist_id	A unique identifier per artist	
genre	Indicates the main genre an artists is associated with	nominal
featured	Indicates whether an artists has been shown on the front page	nominal
monthly_listeners	The average number of listeners per month for an artist	

Table 3 Attributes of artists.csv