

ТЕОРИЈА УЗОРАКА (ЗВ) - Писмени испит
20.8.2025.

1. У датотеци *airlines_flights_data* налазе се подаци о летовима између 6 градова у Индији током периода од 49 дана.
 - а) Изабрати прост случајан узорак без понављања од 3 авио-компаније, па оценити укупан број летова из Делхија (*source_city = Delhi*) за Мумбај (*destination_city = Mumbai*). Одредити оцену дисперзије те оцене.
 - б) Изабрати узорак од 2 авио-компаније, без понављања, са вероватноћама избора пропорционалним броју летова, па оценити просечну цену карте економске класе Хорвиц-Томпсоновом оценом. Одредити оцену дисперзије те оцене.
2. Датотека *spotify_history* садржи комплетну историју пуштања песама од стране једног корисника платформе *Spotify*. У датотеци су дати записи о сваком појединачном пуштању песама са платформе, без обзира на то да ли је песма одслушана до краја, прескочена, да ли је покренута са неке плеј-листе, из претраге, итд.
 - а) Извршити стратификацију према начину пуштања (*reason_start*). Изабрати стратификован случајан узорак без понављања обима 5000 користећи Нејманов избор, па оценити пропорцију прескочених песама (*skipped = TRUE*) и одредити дисперзију те оцене.
 - б) Изабрати систематски узорак обима 127 пуштања песама Лејди Гаге (*artist_name = Lady Gaga*), па оценити укупно време (у минутима) које је корисник провео слушајући наведеног извођача, а затим одредити дисперзију те оцене. У колони *ms_played* дато је време репродукције у милисекундама.