Probability & Statistics: Research project (Interim report)

Taras Levytskyi, Anna Kryvulia, Arsen Botsko

November 2025

1 Team

Taras Levytskyi (Computer Science); Anna Kryvulia (IT & BA); Arsen Botsko (Computer Science)

2 Goal

The goal of our project is to analyze the variables and events that influence the duration of air alarms.
To achieve this, two datasets will be used - one from
`https://github.com/Vadimkin/ukrainian-air-raid-sirens-dataset`
that includes every single air alarm from March 2022 to the present day, and lists their location(oblast, raion, hromada), location level (e.g. oblast-wide, hromada-wide), start/finish time and source. The other dataset includes alarms and their types (reasons they were declared, e.g. drone/missile attacks, MIG-31K and more) for Lviv oblast' only, from 2023 up to today. It has all alarms listed with the following data: start time, end time, duration (in minutes), danger type (or unknown).

With this data we will test the following hypotheses:

1. For each air alarm type:
   $H_0 : \mu = \mu_{type}$
   $H_1 : \mu \neq \mu_{type}$,
   where μ is the mean of the whole population, and μ_{type} is the sample mean of the type

2. H_0 - start time of an alarm and it's duration are independent
   H_1 - there is a dependence between the time of the start of an alarm and it's duration

We will also test which factors influence air alarm duration the most. Then our team will make a statistical model that uses those variables as input, and gives a prediction of the duration of an air alarm.