НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ

імені Ігоря Сікорського»

Факультет прикладної математики

Кафедра прикладної математики

Звіт

з лабораторної роботи №*1*

з дисципліни «Аналіз Данни»

на тему:

*Аналіз EDA*

|  |  |
| --- | --- |
| Виконав: | Керівник: |
| студент групи КМ-01 |  |
| *Матюшко П. В.* |  |

Київ— 2023

ЗМІСТ

[Ідея роботи 2](#_Toc94713306)

[Мотивація проведення досліджень 3](#_Toc94713307)

[Опис данних 3](#_Toc94713308)

[Результати EDA 5](#_Toc94713309)

[Висновки 6](#_Toc94713310)

[Додаток 17](#_Toc94713311)

# Ідея Роботи

# Мотивація проведення дослідження

- досліднецькі питання .

- гіпотези, очікування

# Опис данних та кроків із їх підготовки

- походження данних

Kaggle

- що було виконано для очищення данних і в який спосіб

**Стовпчики, які були видалені:**

telecomHybrydUploadSpeed - 80% відсутніх данних

Housenumber - не має дослідноцької цінності

geo\_krs - у нас есть район

thermalChar - не понятно что это

regio3 - у нас є regio2,regio1

telecomHybryd - 83%, відсутні

geo\_bln - same as regio1

Street - не має дослідноцької цінності (чи можливо поріівнювати вулиці в одному й тому самому місті)

Streeplain - теж саме що й назва вулиці

livingSpaceRange - не зрозуміло шо це

description - тест опис кв, який складно опрацювати(або можно використати пошук слів)

Facilities- опис кв додатковий

electricityBasePrice - куча НА или одно и тоже значение

elecrtycytiesperKwh - не має дослідноцької цінності

**Стовпчики, з помилковим типом данних :**

heatingType - chr -> Faktor

telekomTvOffer -chr -> Faktor

firingTypes - chr -> Faktor

typeOfFlat - chr -> Faktor

**Стовпчики, з відносно великою кількостью NA :**

serviceCharge - замінити середнім

heatingType - замінить найпощириншим типом

telekomTvOffer - замінить найпощириншим типом

yearConstructed - видалити?

totalRent - середнє чи згрупувати по якимость фітчам і взяти середнє

noParkSpaces - хз що робить

firingTypes - замінить найпощириншим типом

houseNumber - хз

condition - хз

petsAllowed - хз

typeOfFlat - хз

Floor - хз

numberOfFloors - хз

heatingCosts - замінити середнім

energyEfficientyClass - нічого не робити

lastRefubish - нічого рне робити

- основні характеристика очищенних данних

# Результати EDA

# Висновки

# Додаток