KEEPING OUR ROADS SAFE

TRAFFIC ACCIDENT ANALYSIS - THAILAND

MCROSOFT CODE WITHOUT BARRIERS HACKATHON 2022

Prepared by
Ratri Maria

ABOUT MYSELF

An enthusiast of AI/Machine Learning and a proud Anthropologist.

More than ten years of experience in various disciplines, including marketing, consumer insights, communication, data analytics, and user experience.

I am currently developing Changemaker Asia, a platform dedicated to solving underserved, fragmented, and siloed ecosystems of social impacts in Asia.

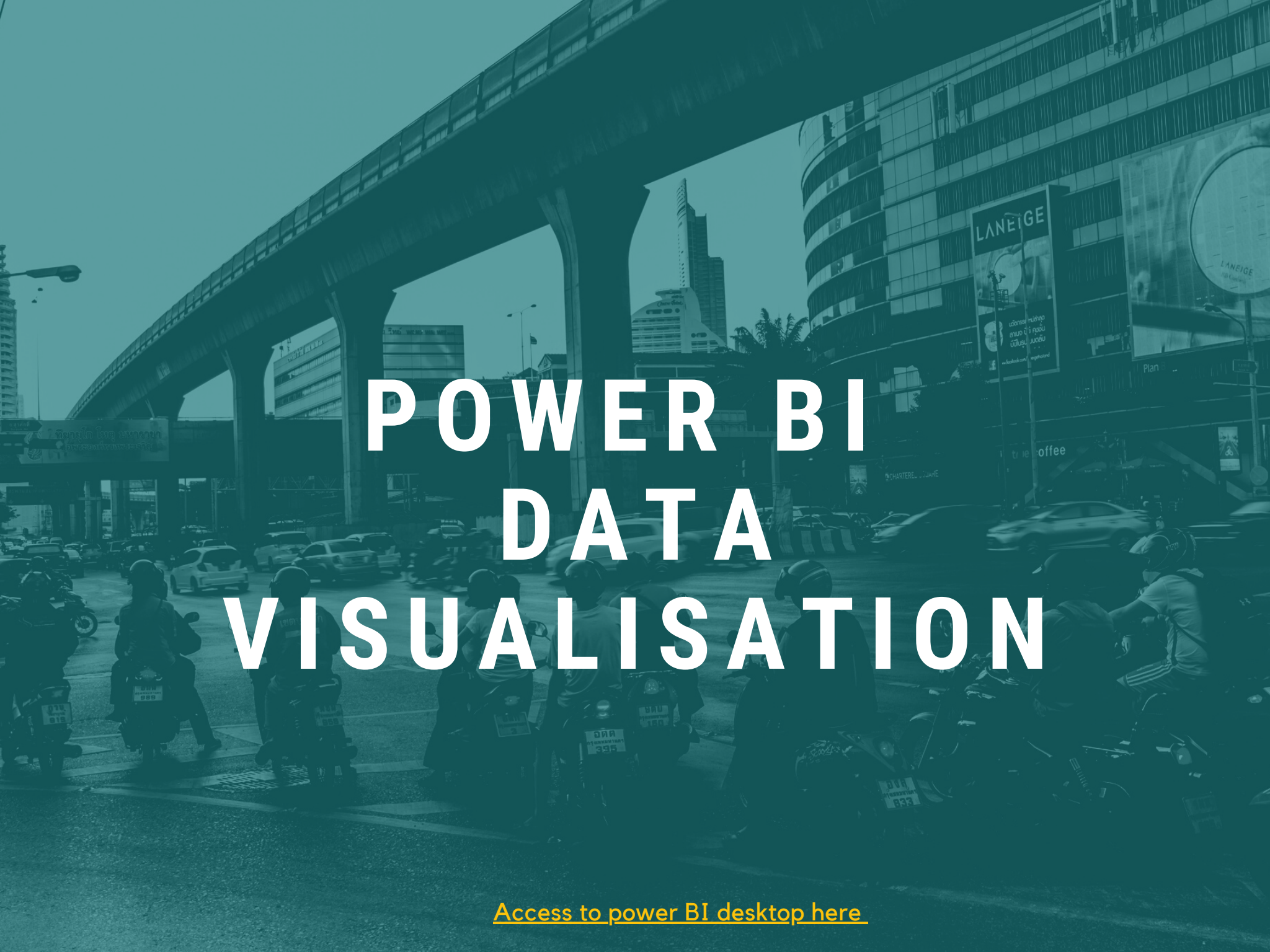
PROBLEM STATEMENT

Thailand has a mission to reduce car accidents and improve road safety. Therefore, it is crucial to understand the causes of the accidents and identify risky areas.

Participants are expected to perform two following tasks:

1. Data visualization to analyze possible factors of the accidents. It is highly recommended to use other open datasets (for example, datasets from https://opendata.data.go.th/dataset) to support your analysis.

2. A machine learning model to predict car accidents in each area and suggest how to prevent car accidents in the future.

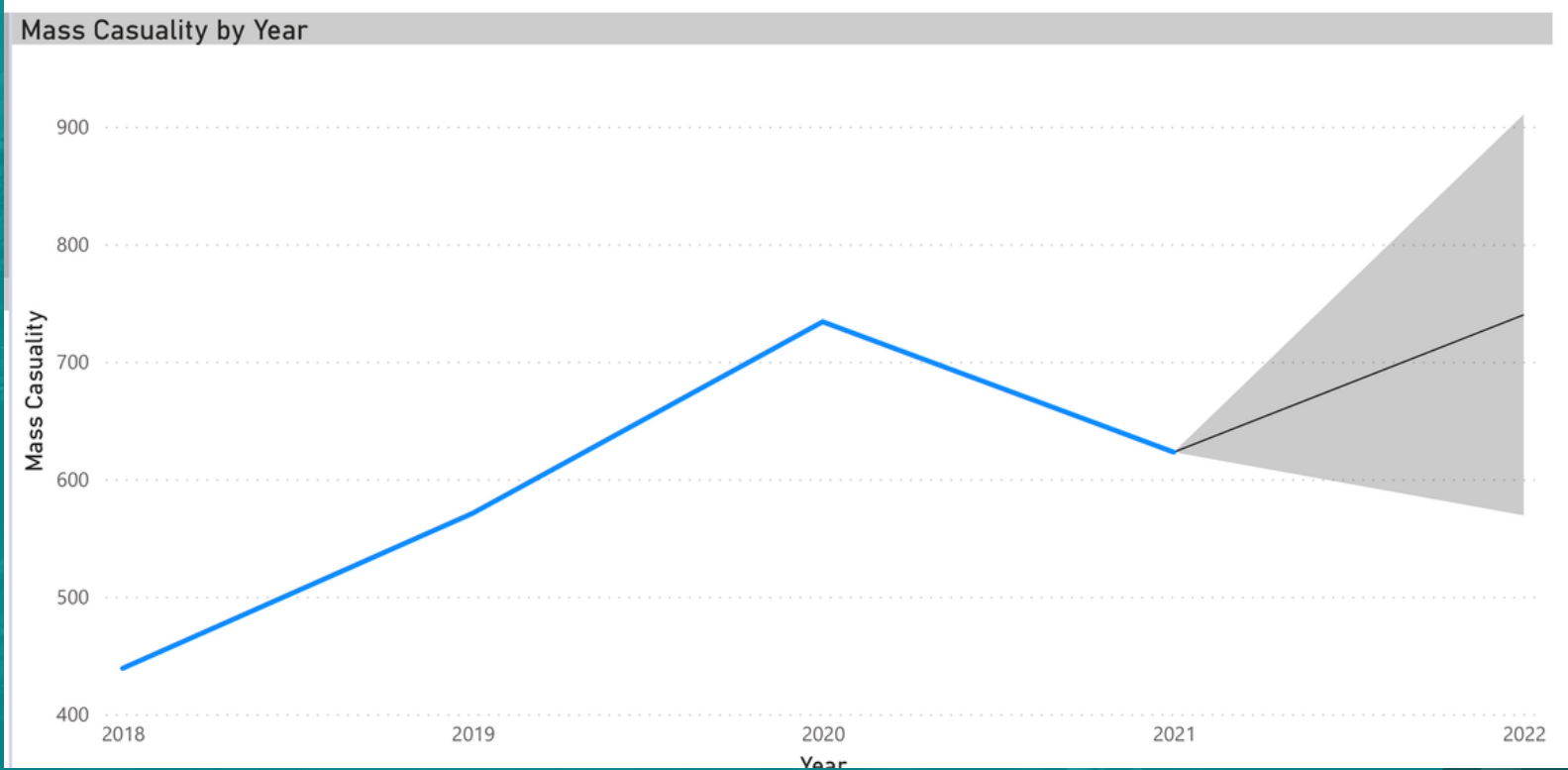


POWER BI DATA VISUALISATION

Access to power BI desktop here

POWER BI

PROJECTING THE FUTURE TRENDS IN MASS CAUSALITY OVER THE NEXT 12 MONTHS

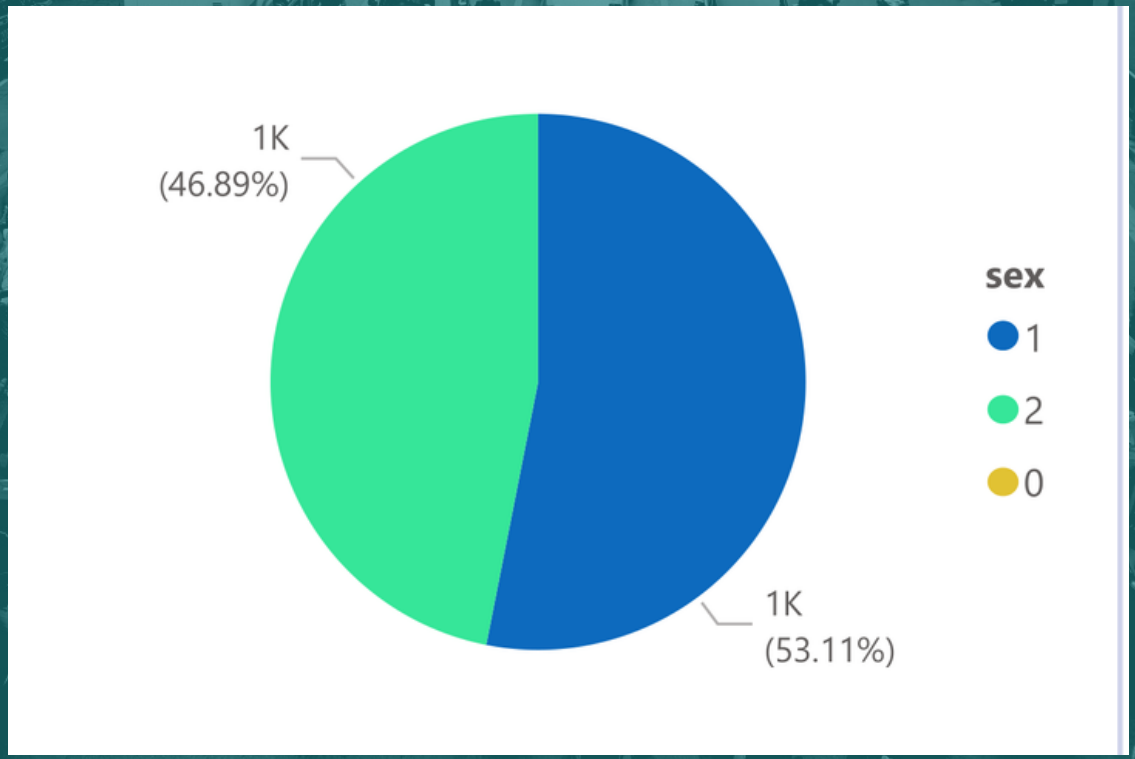


Access to power BI desktop here

POWER BI

NUMBER OF INJURED PERSON, BY GENDER

Overall, men are more impacted by traffic accidents*

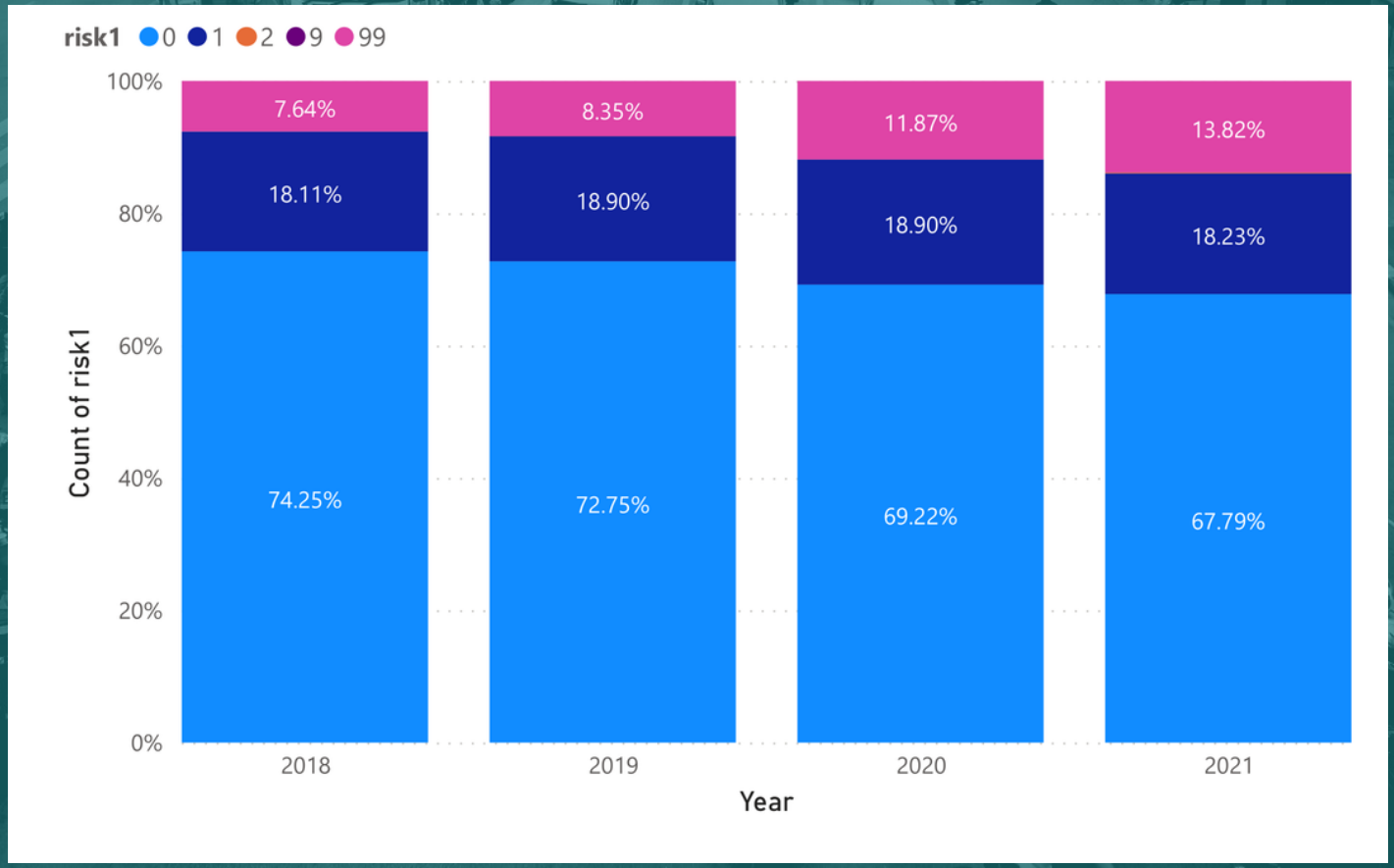


*Phase 2 : to cross check with Thailand Population Distribution - Open Data

POWER BI

THE INJURED PERSON DRUNK ALCOHOL (RISK 1) BY YEARS*

1 in 5 of the injured person across the year were found consumed alchohol. The trend has stayed the same from 2018 to 2021.

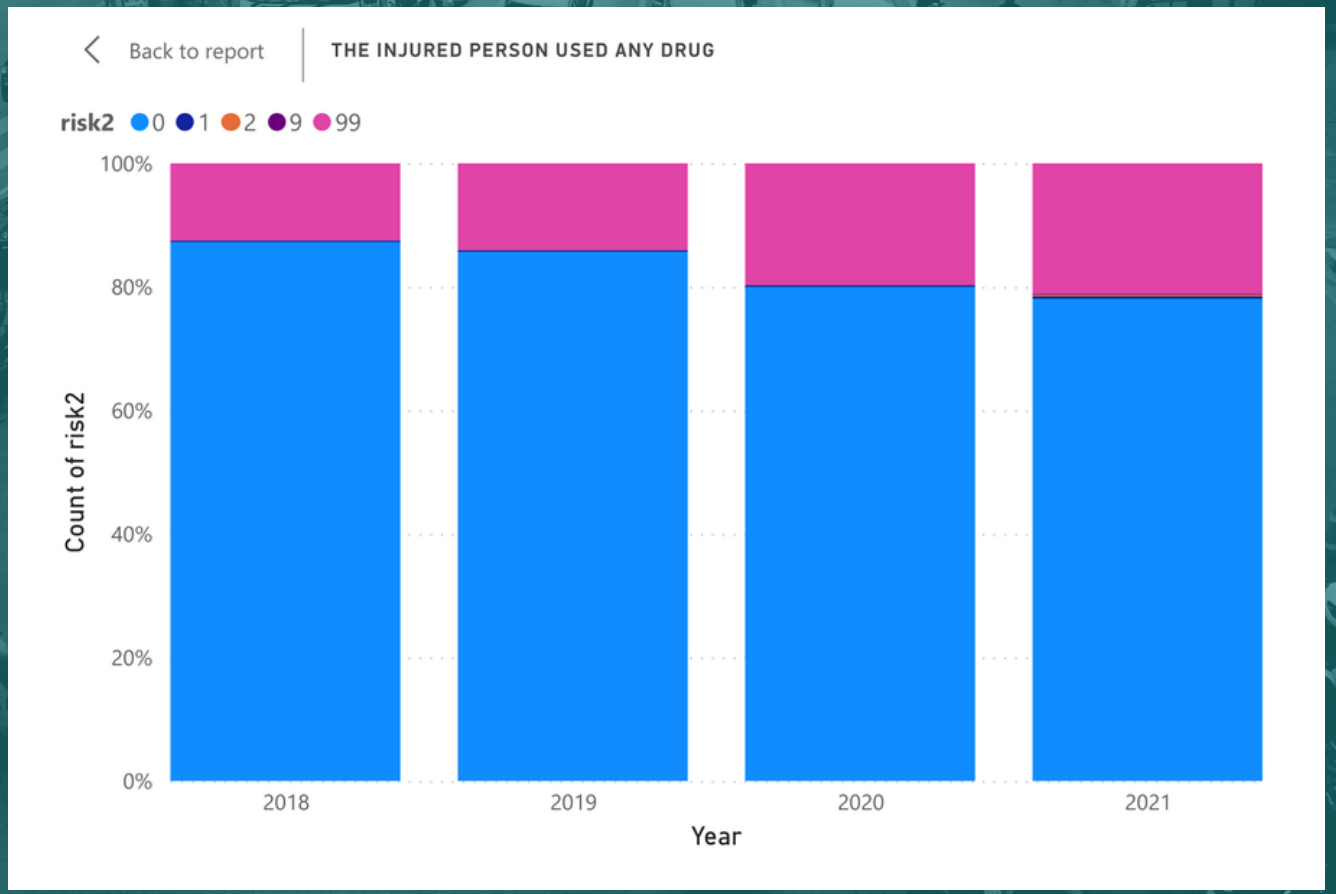


*Phase 2 :To slice the data by gender and age

POWER BI

THE INJURED PERSON USED ANY DRUG (RISK 2) BY YEARS

In risk 2, there is an increased trend of missing data. Underreporting in accident records is expected, as injured individuals may be reluctant to self-report.



*Phase 2 :To slice the data by gender , age and day/time of accident (weekdays vs weekends vs public holiday)

POWER BI

THE INJURED PERSON FASTEN SEAT BELT(RISK 3) BY YEARS

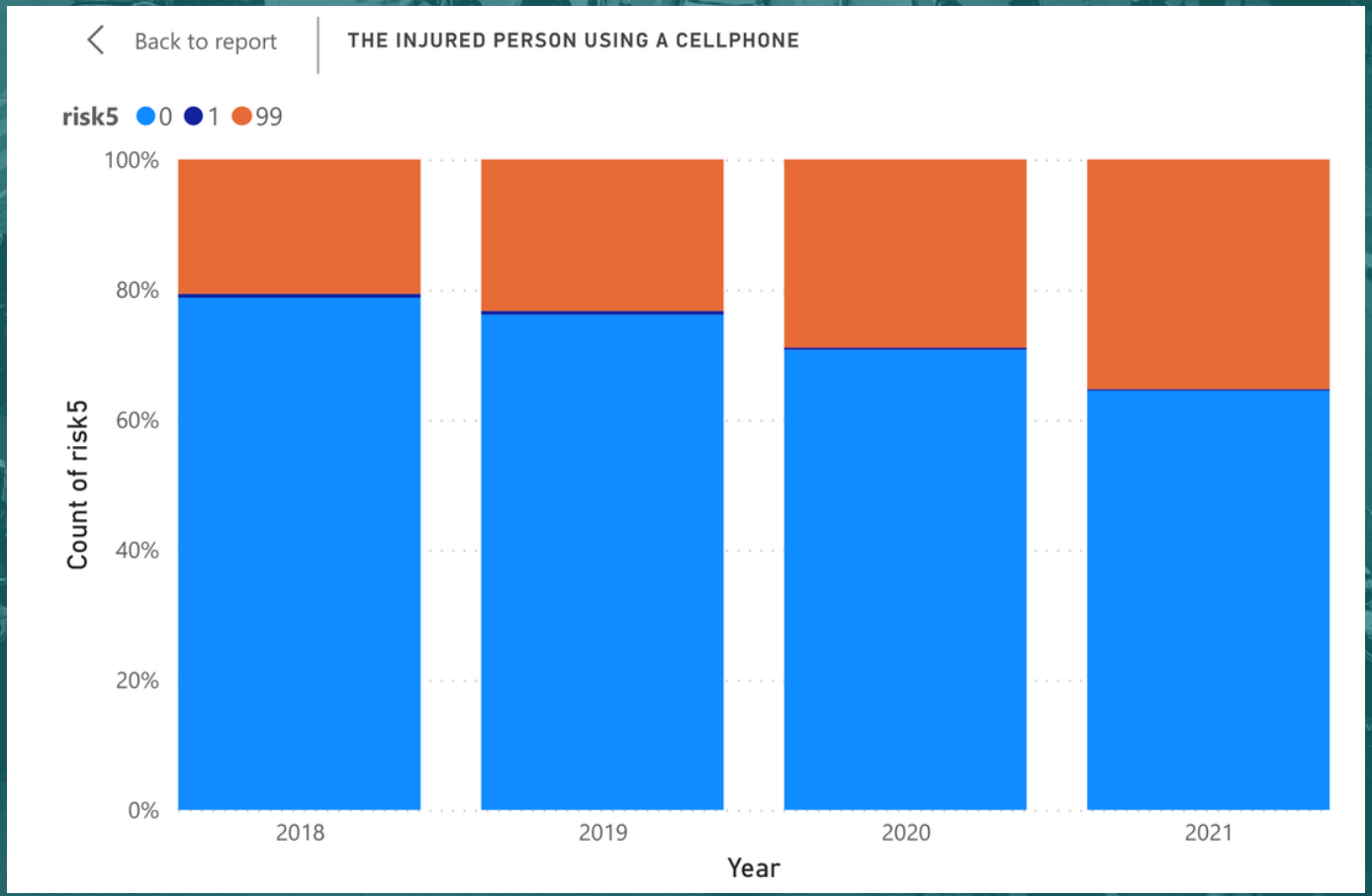
The number of injured people who do not fasten their seatbelt is on the rise.

*Phase 2: Split the data by type of injured person (driver vs passenger) and type of vehicle (public vs private transports). Also to see if there's a statistically significant trend difference between rural vs urban areas related to compliance with the seat belt regulations between the front seats than in the rear seats (potential source of data : gross provincial product, level of literacy and law penalty and its association with seat belt use.

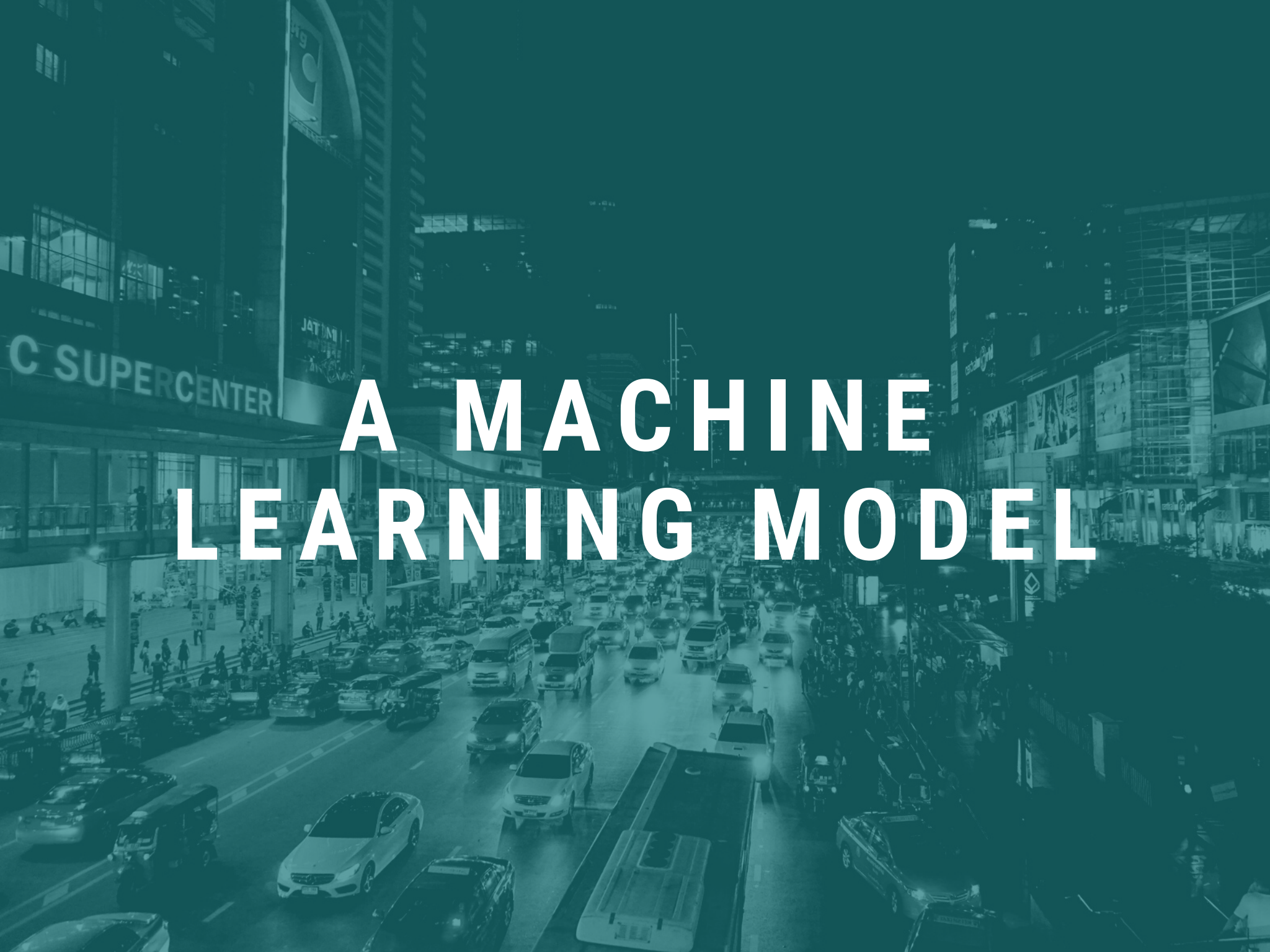
POWER BI

THE INJURED PERSON USED MOBILE PHONE (RISK 4) BY YEARS

Although the data indicate a decrease in the use of mobile phones, there is a trend of absent data in risk 4; further analysis is required to identify the cause of the problem

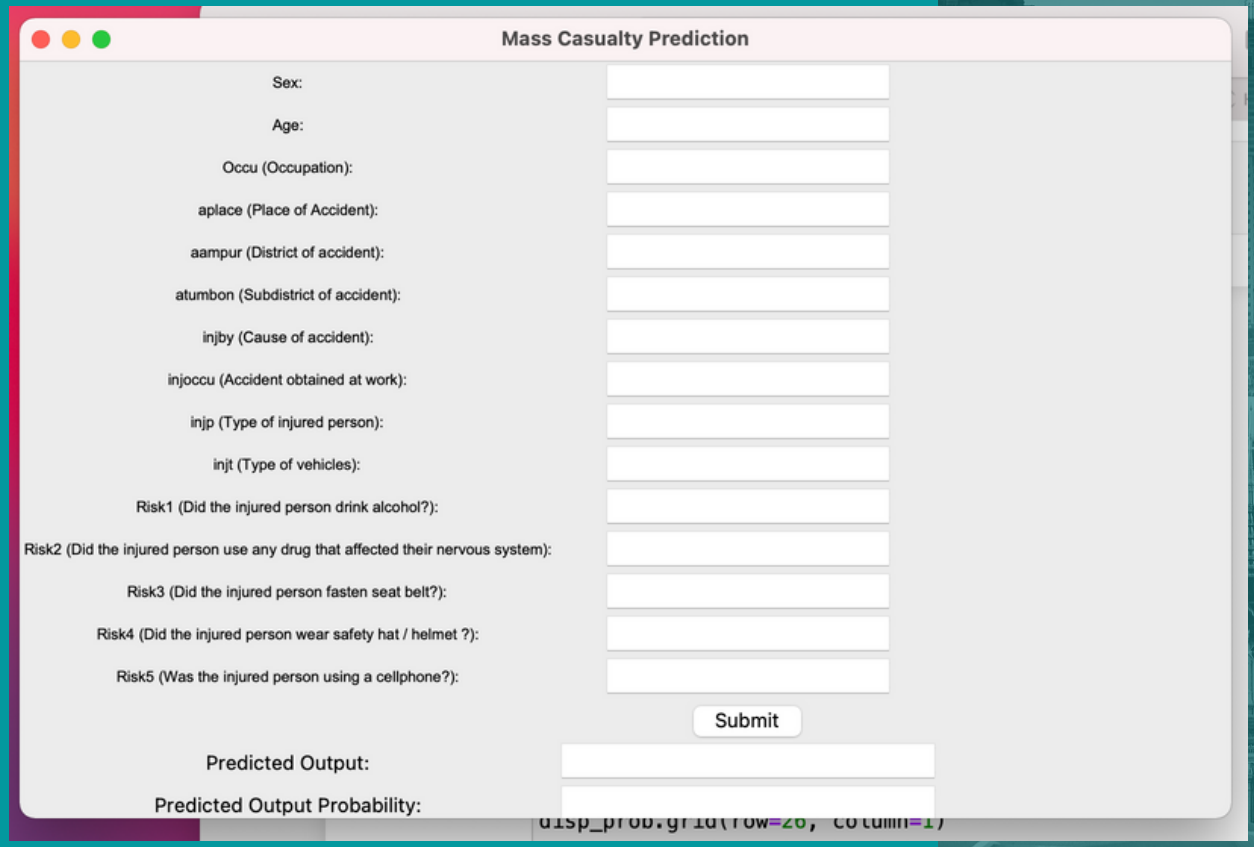


*Phase 2 :To slice the data by gender , age and day/time of accident (weekdays vs weekends vs public holiday)



A MACHINE LEARNING MODEL

MASS CASUALITY PREDICTION



Because the web service deployed in Azure couldn't be connected to Power BI,
I built a simple interface for data entry and predicted output probability.

You can play the video by clicking the image or here to view the video on dropbox

LIMITATIONS & FUTURE WORK

- This current study faced several constraints, such as limited to only four databases from Injury surveillance system owned by Division of Injury Prevention, Department of Disease Control, Ministry of Public Health, Thailand. The datasets focus on the injured people that have got treatment in hospitals.

- There is a possibility that other techniques are present which are not listed in this presentation.

- For further research can be supplemented by utilising road type data for model training. This will improve the prediction of traffic accident and further enhance the accuracy of predictive models, addressing the responsible factors for accidents likely to happen.

THANK YOU
TAXI-METER
TAXI-METER