

Projet data-science

2020-2021

Our team



Eya Loukil
Hana Bhar
Nour Soffelgil
Omar Dahmen
Yassine Ayedi

Sommaire

- I. Introduction
- II. Business and Data comprehension
- III. Data Preparation
- IV. Data modeling and Evaluation
- V. Data Visualization
- VI. Conclusion and perspectives

I. Introduction

Who is our client?

It's a B2B segmentation



+



I. Introduction

The problematic



Accidents



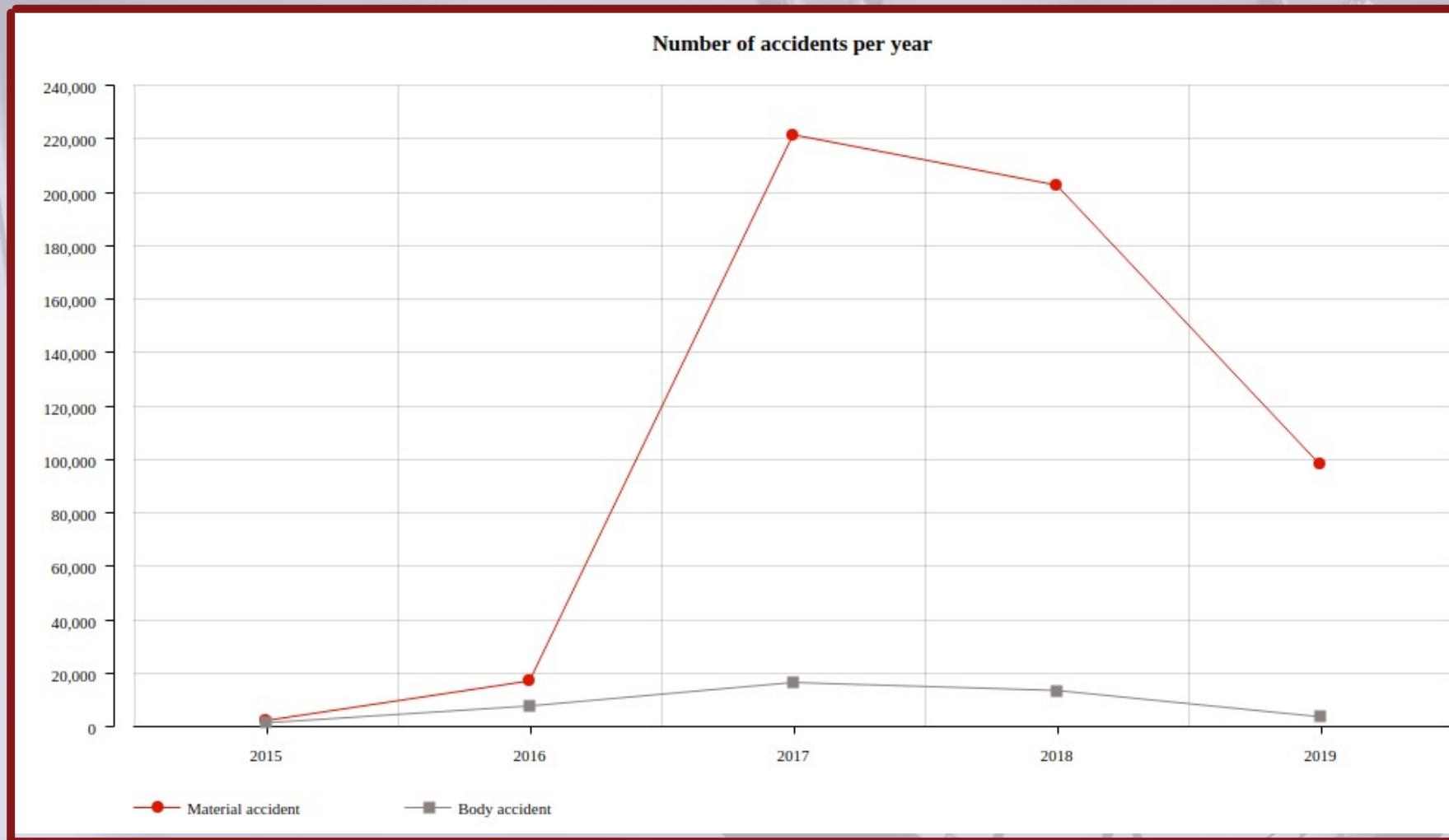
Frauds increase



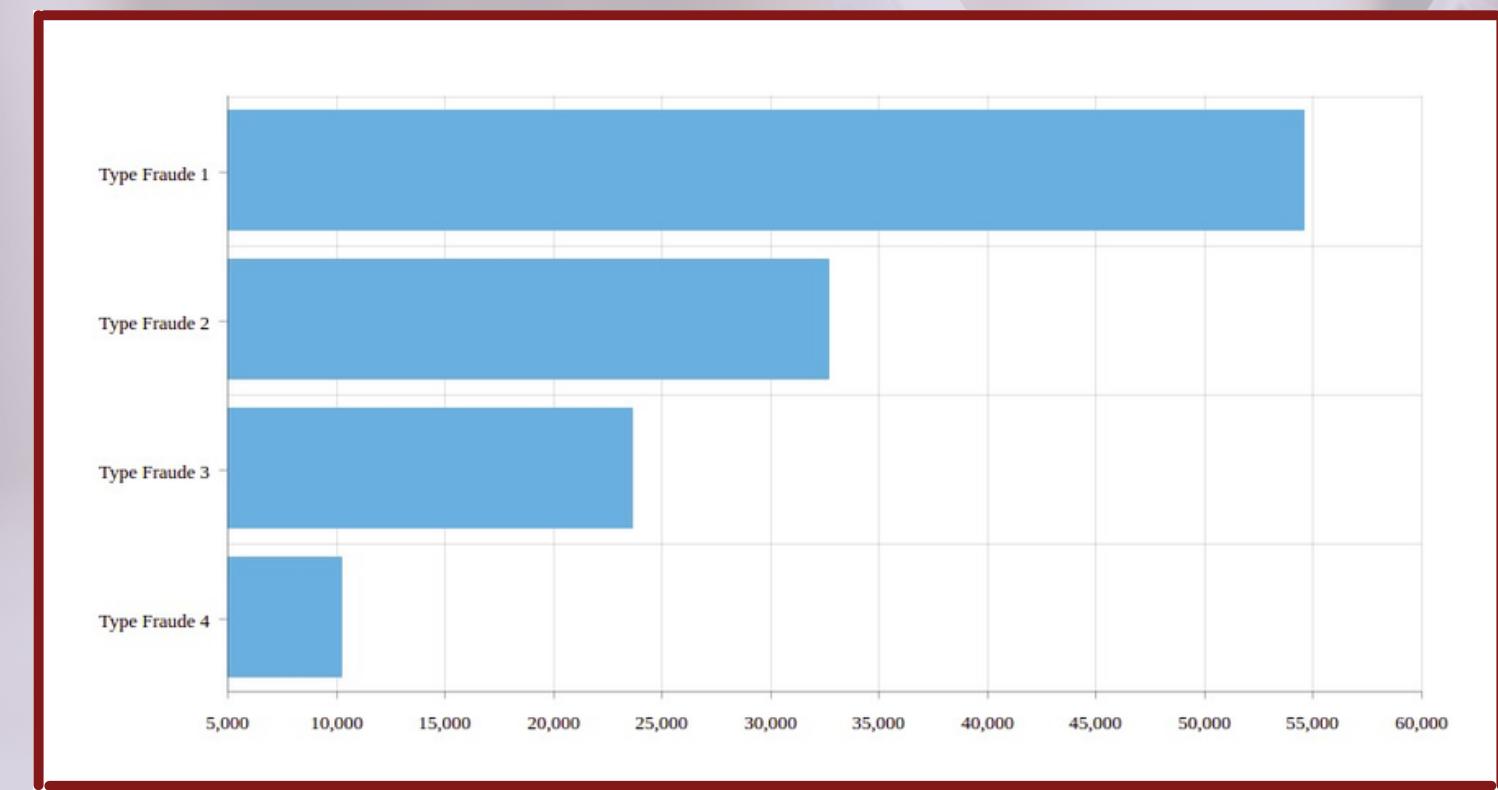
Unreliability decrease

I. Introduction

Some statistics



Number of accidents distribution per year



A bar chart illustrating the different types of fraud

I. Introduction Solution

Our product



Accident's predition



Our services



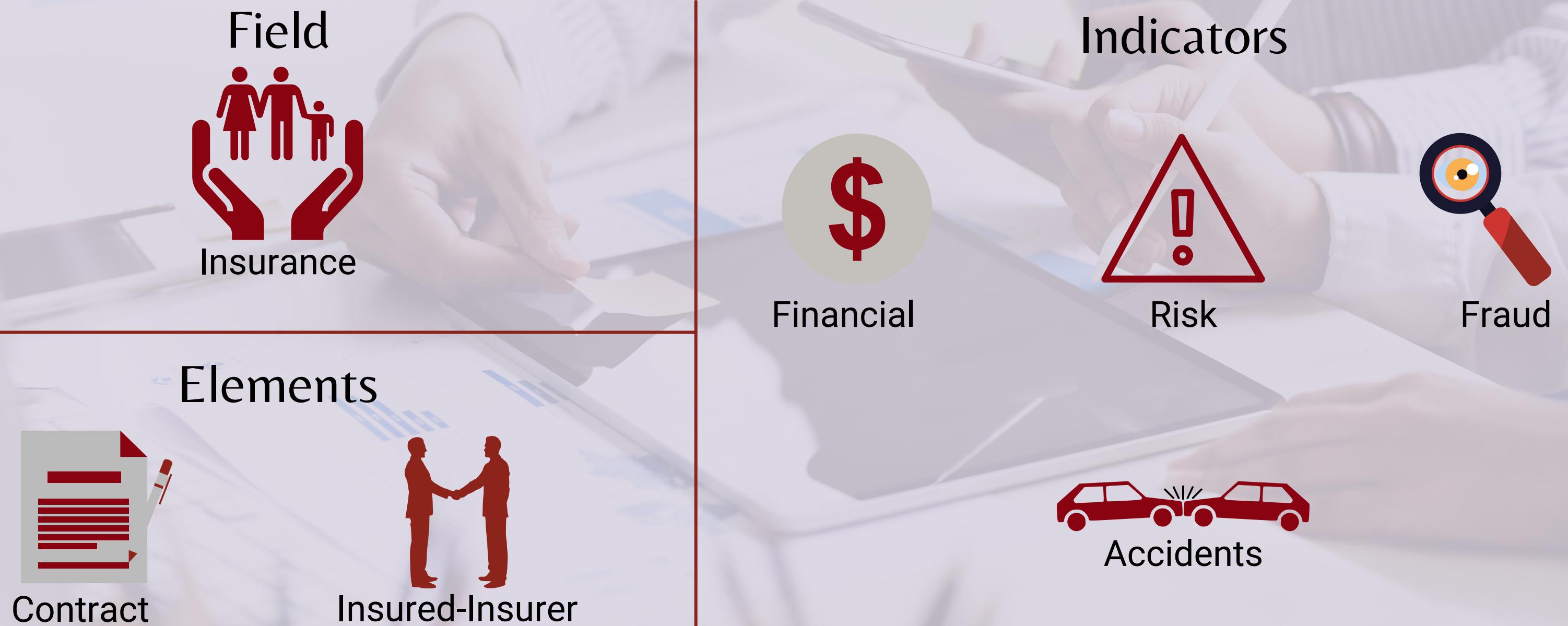
Bonus-malus
class



Fraud detection

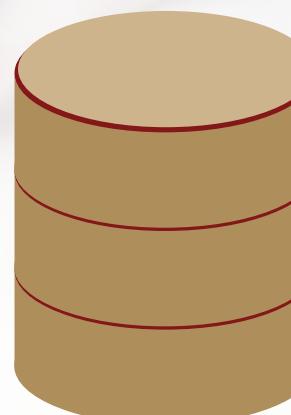
II. Business and Data comprehension

1. Business comprehension

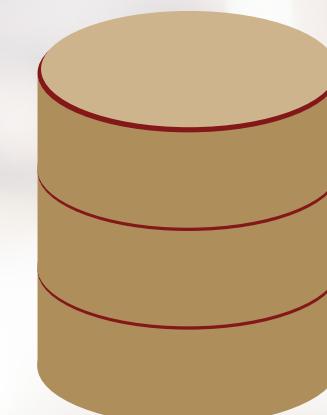


II. Business and Data comprehension

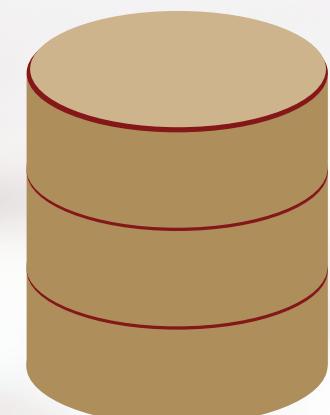
2. Data comprehension



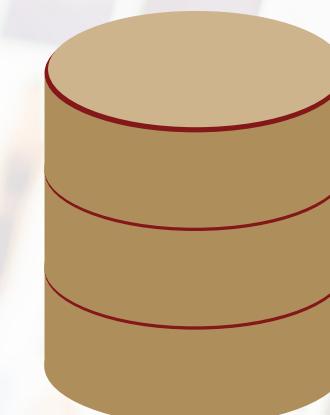
Vehicle



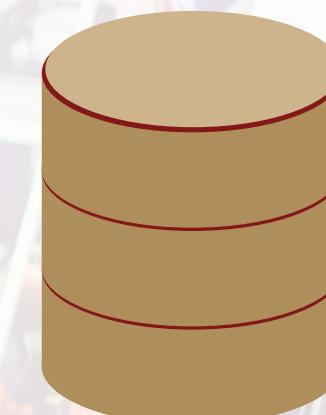
ClassBonusMalus



Sinistres



Assure

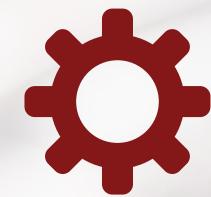


Police

II. Business and Data comprehension

2. Data comprehension

Key data
"Vehicule"



Car use and fiscal power



Car groups



Energy

Why? To segment the car's category.

II. Business and Data comprehension

2. Data comprehension

Key data

"Police"



Police's

Type
Nature
State
Intermediate type

Why?

To validate the police

II. Business and Data comprehension

2. Data comprehension

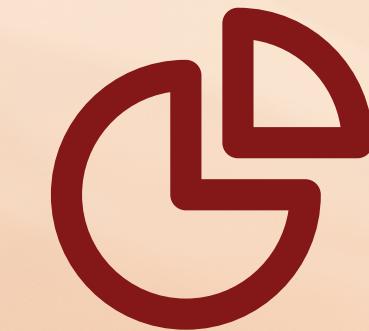
The result



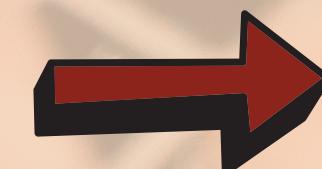
Driving experience



Validate police



Car's category

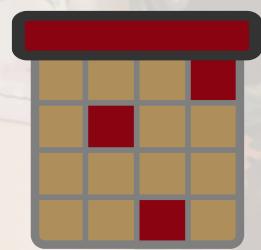


Service: bonus malus class prediction

II. Business and Data comprehension

2. Data comprehension

Key data
"Sinistre"



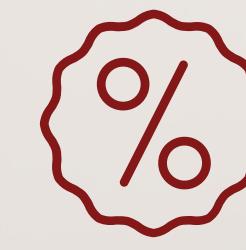
Sinistre date



Sinistre hour



Sinistre location



Percentage of responsibility

Why?

Sinistre's features

Why?

To validate the credibility
of the insured

II. Business and Data comprehension

2. Data comprehension

The result



Sinistre's features



Percentage
of responsibility



Validate police



Car's category



Class Bonus-Malus



Service: Fraud detection

II. Business and Data comprehension

2. Data comprehension

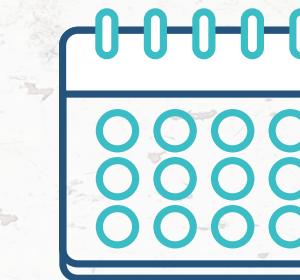
To prepare



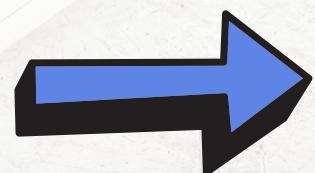
Weather state



Accident location



Accident date



Service: Estimated number of accidents

III. Data preparation

1. Internal data

Deleted



Empty
columns

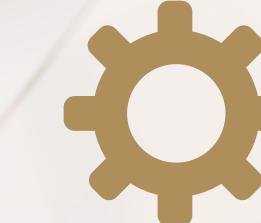


Nearly
columns

Merge



Vehicule



Usage vehicule



Police



Assuré

Marque vehicule

III. Data preparation

1. Internal data

The result



III. Data preparation

1. Internal data

a. Bonus-malus class table

Delete



CodeAssureId
Dateaffectation
Etatvehicule
Etatvehicule
DateAjout
etc

Calculate



Driving experience

Division



Cars brands into group

Encodage



Categorical variables

III. Data preparation

1. Internal data

a. Sinistres table

Fraud detection

Type 1

Insured have accidents more than 5
and % of responsibility>50%

Type 2

Insured % of responsibility>100%

Type 3.

One police has multiple car

Type 4

Type of police is flotte and number
of cars >5

Type 5.

Inadequat bonus-malus class with
the coefficient



III. Data preparation

2. External data

Web scraping



historique-meteo.tn

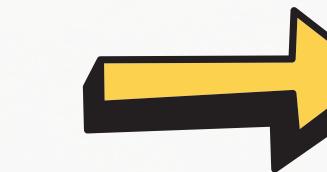


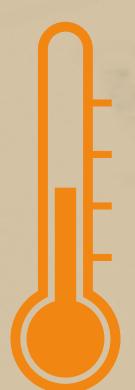
Table weather



III. Data preparation

2. External data

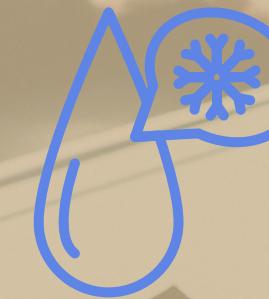
Result



Temperature



Humidity



Precipitations



Wind speed



Visibility

IV. Data modeling and Evaluation

1. Data modeling

Class bonus-malus - Fraud detection - Total accidents prediction

Decision Tree

Advantages

Applies to both categorical and numerical data

Disadvantages

Higher risk of over-fitting

RandomForest

It's considered as a highly accurate and robust method

It's difficult to interpret

LGBM

Better accuracy than any other boosting algorithm

Less documentation available.

XGB

Versatile

Higher risk of over-fitting

CatBoost

An innovative algorithm for processing categorical features

Difficulty in tuning the parameters

IV. Data modeling and Evaluation

1. Data modeling

Bonus Malus Class Classification

We use 11 bonus-malus classes

Application of undersampling

Oversampling smote 2 classes

Division of class into three groups

Oversampling smote 3 classes

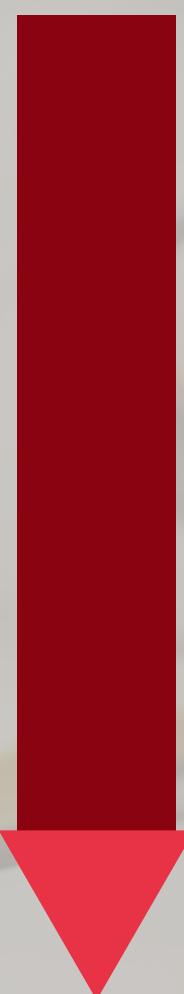
Final model

IV. Data modeling and Evaluation

1. Data modeling

Fraud detection -Total accidents prediction

Application of the same algorithms

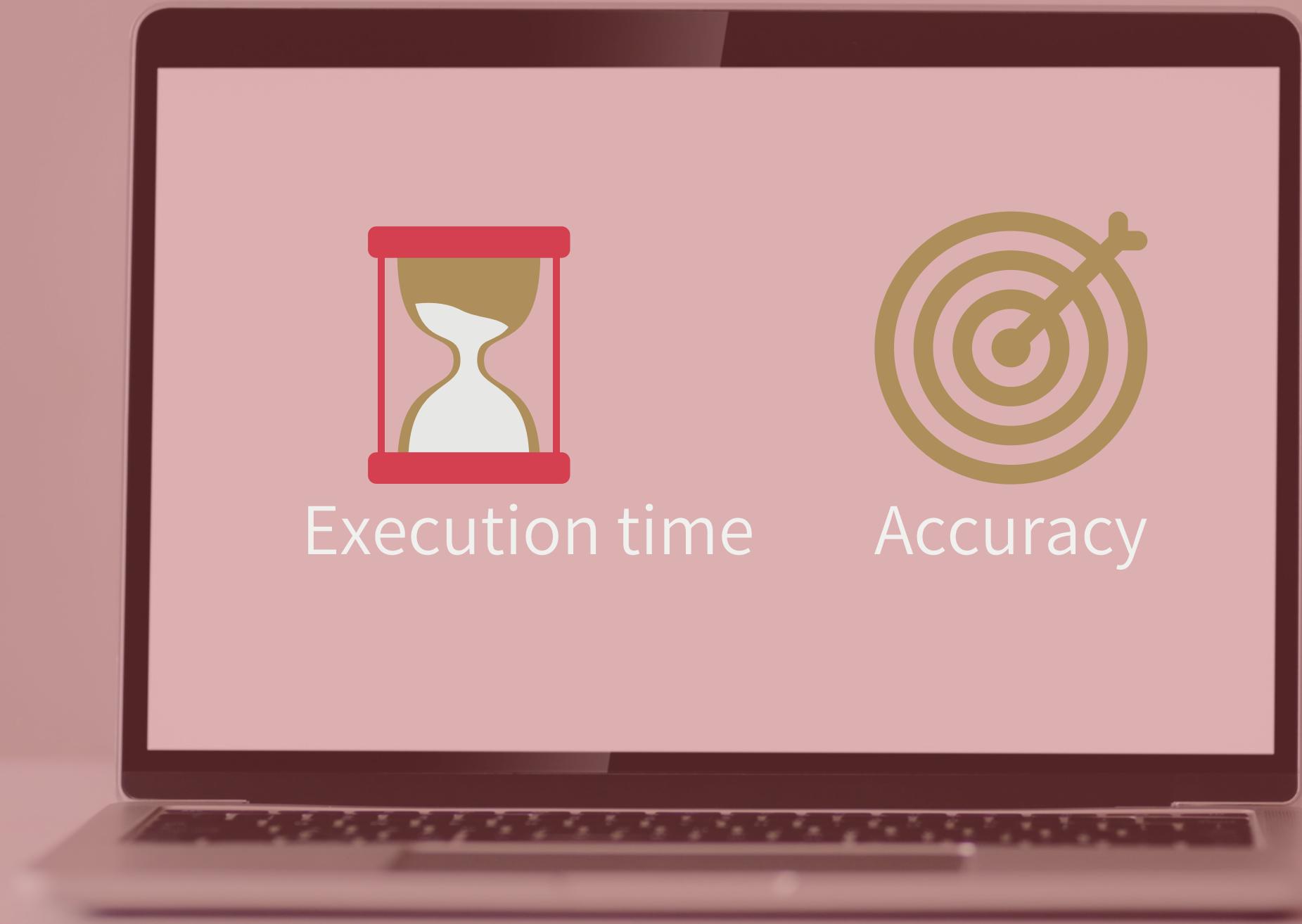


Final model

IV. Data modeling and Evaluation

2. Data evaluation

Selection criteria



IV. Data modeling and Evaluation

2. Data evaluation

Class bonus-malus prediction

1

RandomForest

2

DecisionTree

3

CatBoost

IV. Data modeling and Evaluation

2. Data evaluation

Fraud detection

1

LGBMClassifier

Total accidents prediction

1

LGBMRegressor

V. Deployment and data visualization

1. Deployment

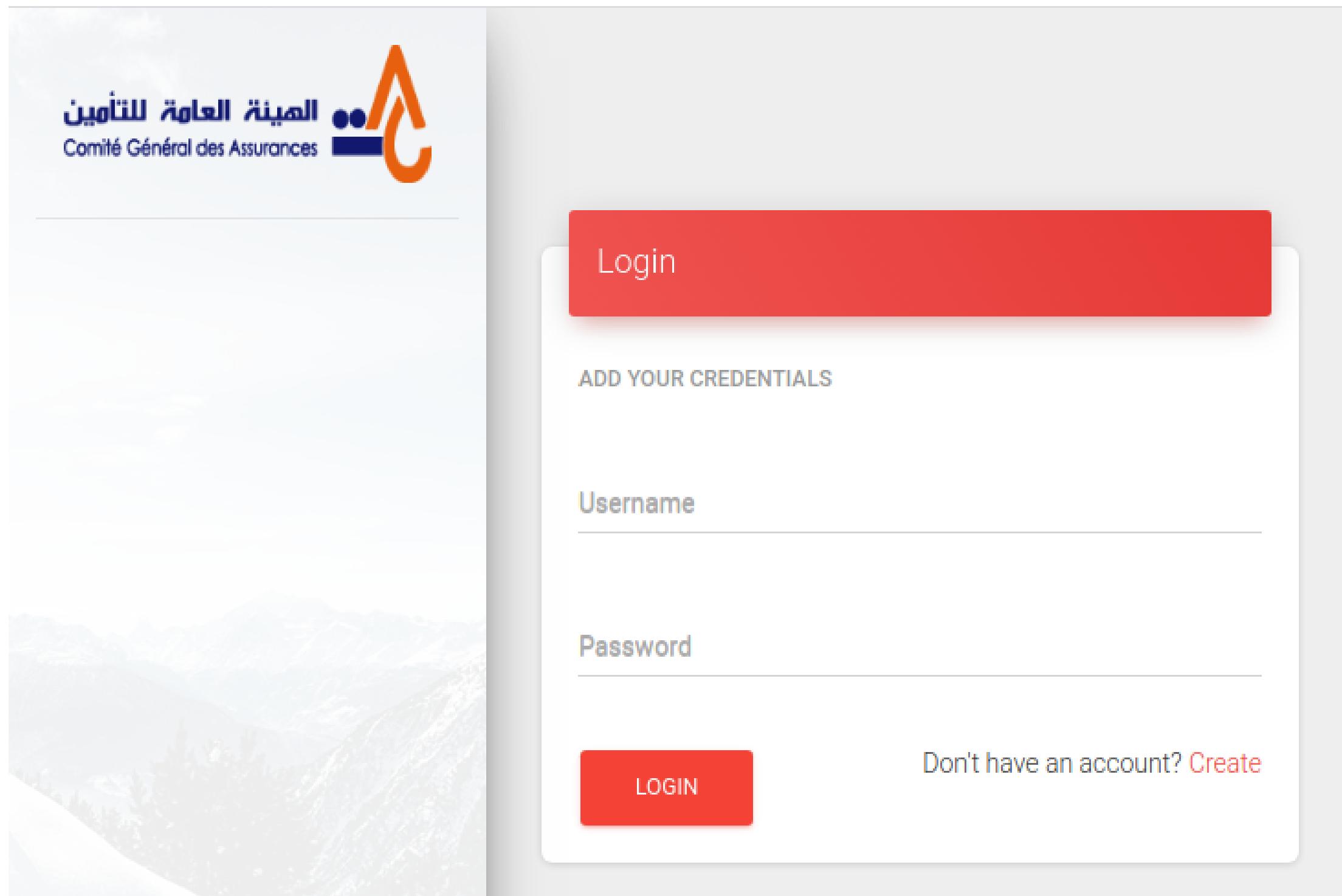
How?

django

Why?

To have a final output

Login interface



Presentation

المهيئة العامة للتأمين
Comité Général des Assurances

Presentation

Dashboard

Bonus Malus Class

Accidents/Weather

Fraud Prediction

Logout

Projet data-science
2020-2021

Dashboard

المدينة العامة للتأمين Comité Général des Assurances

Presentation

Dashboard

Bonus Malus Class

Accidents/Weather

Fraud Prediction

Logout

Search...

Filters

Accidents

Choose the location:

Sinistre location	Total sinistre
tunis	77553
sfax	20623
ariana	17108
sousse	15174
ben arous	13811
nabeul	10597
monastir	7064
gabes	4308
medenine	3605
bizerte	3404
manouba	2927
mahdia	2560
kairouan	2128
beja	1734
gafsa	1471
kef	1319
zaghouan	1004
jendouba	878
Total	190275

Sinistre location: All

Total sinistre by Sinistre location

TUNISIA

ariana
Sinistre location
190K
Totals sinistre

The weather

10.0 Visibility

10.0 Humidity

Total accident by Season

Season	Total accident
automne	~180K
ete	~180K
hiver	~180K
printemps	~180K

Insurances companies

Insureds

Bonus-malus classification

Frauds

Accidents

Class bonus-malus detection

The screenshot shows a web-based application interface for 'Bonus Malus Class' prediction. The top navigation bar includes the logo of 'المدينة العامة للتأمين Comité Général des Assurances' and a search bar. The main content area has a red header 'Prediction Bonus Malus Class'. On the left sidebar, under 'Presentation' and 'Dashboard', there are links for 'Bonus Malus Class' (which is highlighted in red), 'Accidents/Weather', 'Fraud Prediction', and 'Logout'. The main form is divided into two columns: 'Car use' and 'Car groups'. The 'Car use' section contains fields for 'Vehicle Registration Type' (with 'Fiscal Power' as a sub-section) and 'Nature Police' (with 'Renewable' and 'Temporary' options). The 'Police Type' section contains 'Flotte' and 'Individual' options. The 'Car groups' section contains 'Energy' and 'Driving Experience' sections, with 'Police State' (options: Vigor, Terminated, Suspended) and 'Intermediate Type' (options: Agent, Office, Broker) listed. A large red 'PREDICT' button is located at the bottom right of the form.

Accidents/ Weather

Search... 🔍 GRID NOTIFICATIONS USER

Prediction

Accidents / Weather state

Temperature

wind speed

Humidity

Visibility

Precipitation

Date
mm/dd/yyyy

Location

Weather state

PREDICT

Comité Général des Assurances

- Presentation
- Dashboard
- Bonus Malus Class
- Accidents/Weather
- Fraud Prediction
- Logout

Fraud detection

The screenshot shows a web-based application for fraud detection. The left sidebar contains navigation links: Presentation, Dashboard, Bonus Malus Class, Accidents/Weather, Fraud Prediction (which is highlighted in red), and Logout. The main content area has a red header bar with the text "Prediction Fraud". Below the header are several input fields and dropdown menus. On the left, there are fields for "Sinistre date" (mm/dd/yyyy) and "Policy effective date" (mm/dd/yyyy). In the center, there are fields for "Percentage of responsibility", "Sinistre location", "Bonus Malus Class", and "Car groups". On the right, there are fields for "Sinitre hour", "Car use", "Police Type", and "Sinistre nature". At the bottom right is a red "PREDICT" button. The top right corner of the screen shows a search bar, a grid icon, a notification bell with a red dot, and a user profile icon.

Search...

Comité Général des Assurances

Prediction Fraud

Sinistre date
mm/dd/yyyy

Policy effective date
mm/dd/yyyy

Percentage of responsibility

Sinistre location

Bonus Malus Class

Car groups

Sinitre hour

Car use

Police Type

Sinistre nature

Police State

Nature Police

Police Type

Sinistre nature

Vigor

Renewable

Flotte

Material

Terminated

Temporary

Individual

Suspended

PREDICT

Logout

Fraud Prediction

Bonus Malus Class

Accidents/Weather

Dashboard

Presentation

Comité Général des Assurances

المهيئة العامة للتأمينات

A person is shown from the side, wearing a white shirt, working on a silver laptop. Their hands are visible on the keyboard. In the background, there is a large, colorful, abstract map or visualization on a screen, though it is slightly out of focus.

V. Deployment and data visualization

2. Data visualization

Insurance companies

Choose your company's code:

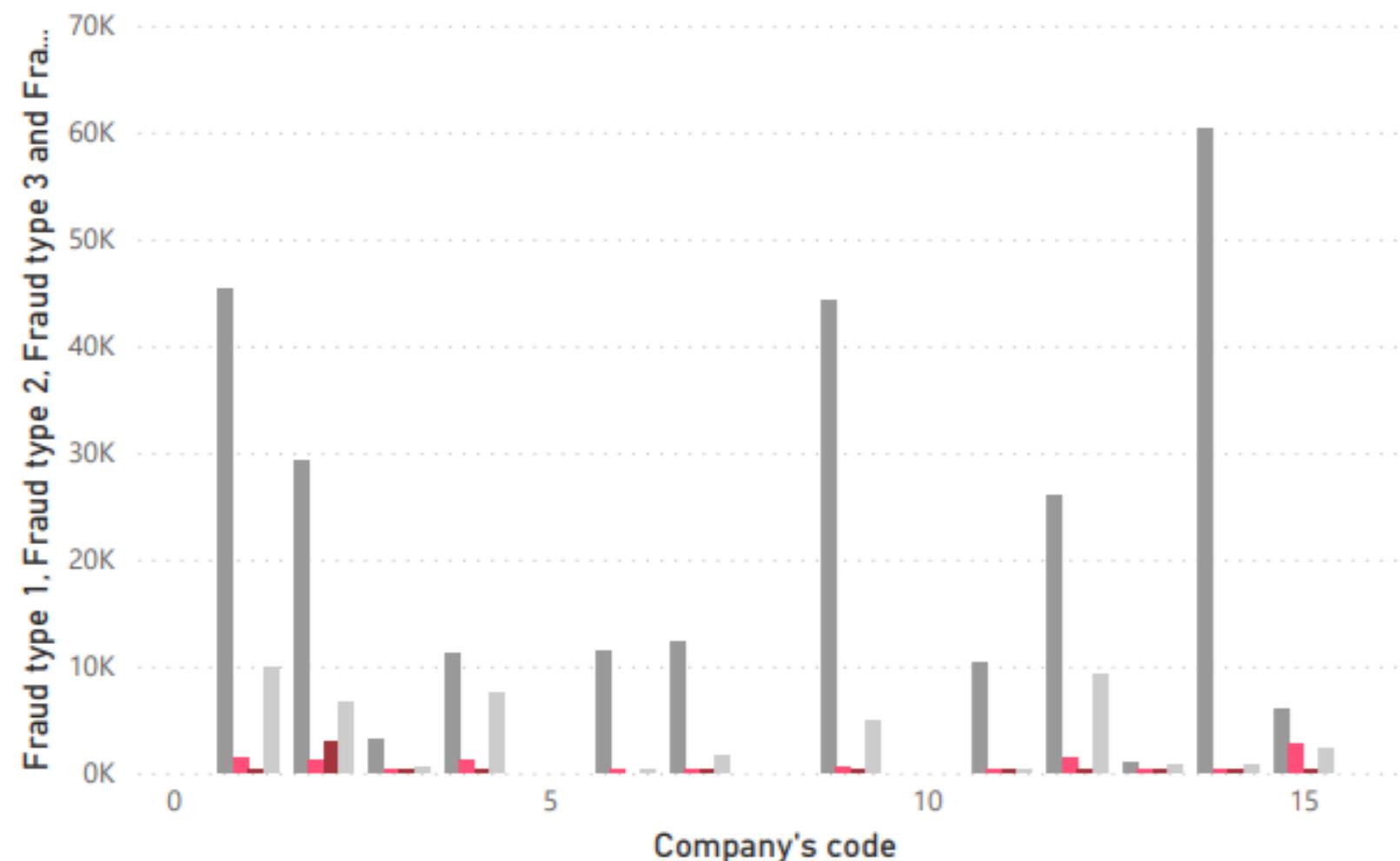
1 15

497.87K

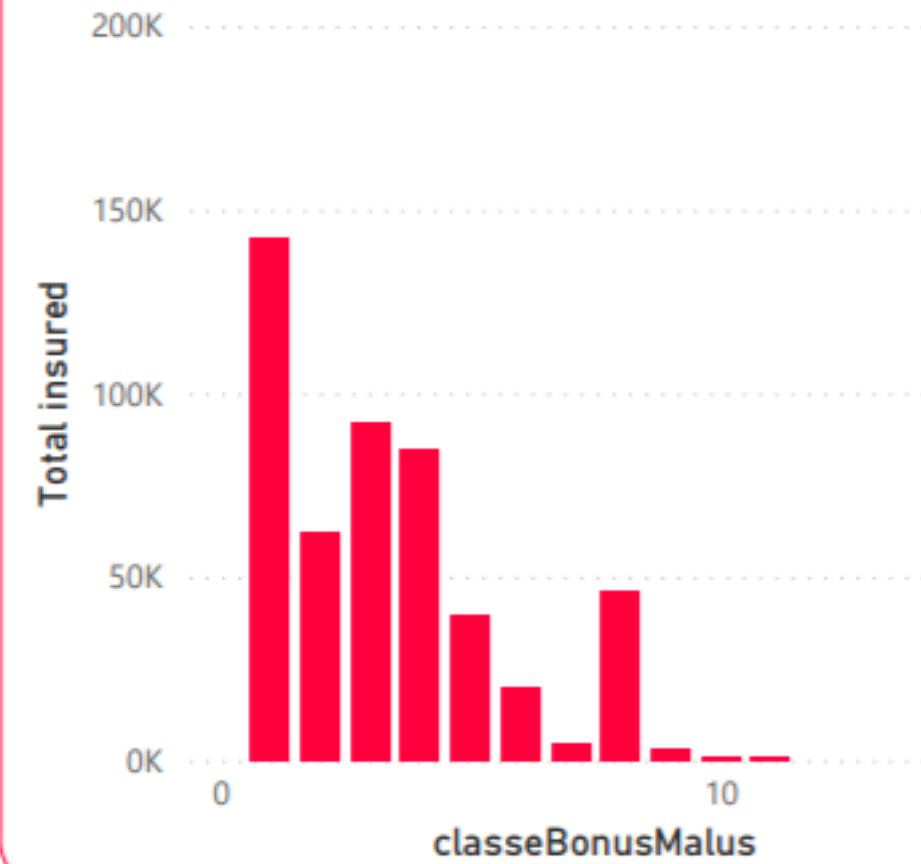
Total insured

Fraud type 1, Fraud type 2, Fraud type 3 and Fraud type 4 by Company's code

● Fraud type 1 ● Fraud type 2 ● Fraud type 3 ● Fraud type 4



Total insured by classeBonusMalus



classeBonusMalus

1
2
3
4
5
6
7
8
9
10
11

Insureds

Bonus-malus class and the coefficient of all insureds

Choose your insured:

481355 1077527



CodeAssure_id	classeBonusMalusCompagnie	coefBonusMalus
481367	1	70
481368	1	80
481374	1	80
481377	1	80
481379	1	80
481382	1	80
481385	1	80
481387	1	70
481388	1	70
481390	1	80
481391	1	80
481394	1	70
481395	1	70
481396	1	80
481397	1	80
481209	1	80
Total	730651	24029500

172bn

CodeSouscripteur_id

731K

classeBonusMalus

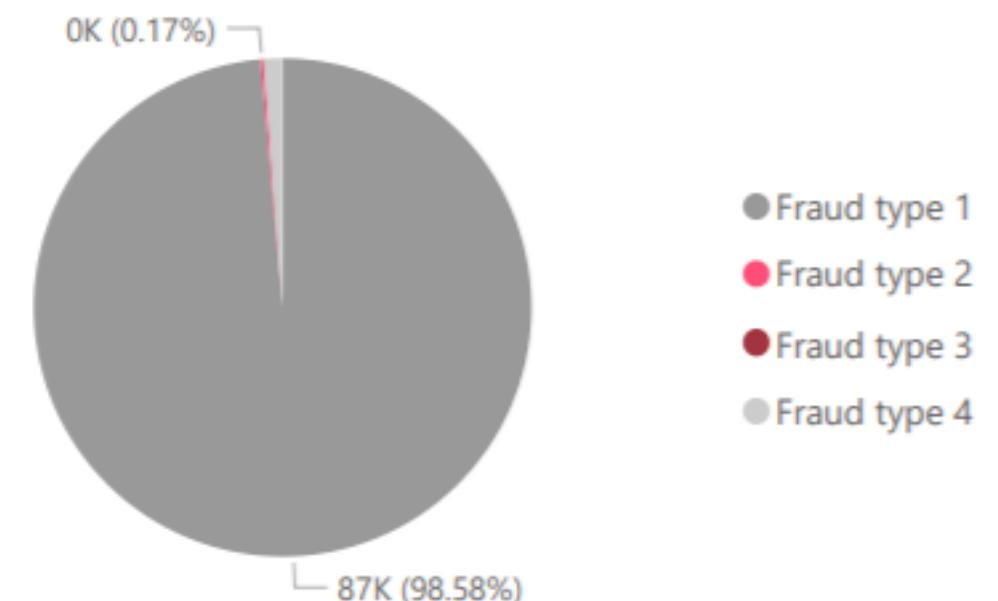
F

Type of police

24M

coefBonusMalus

Fraud type 1, Fraud type 2, Fraud type 3 and Fraud type 4



Bonus-malus classification

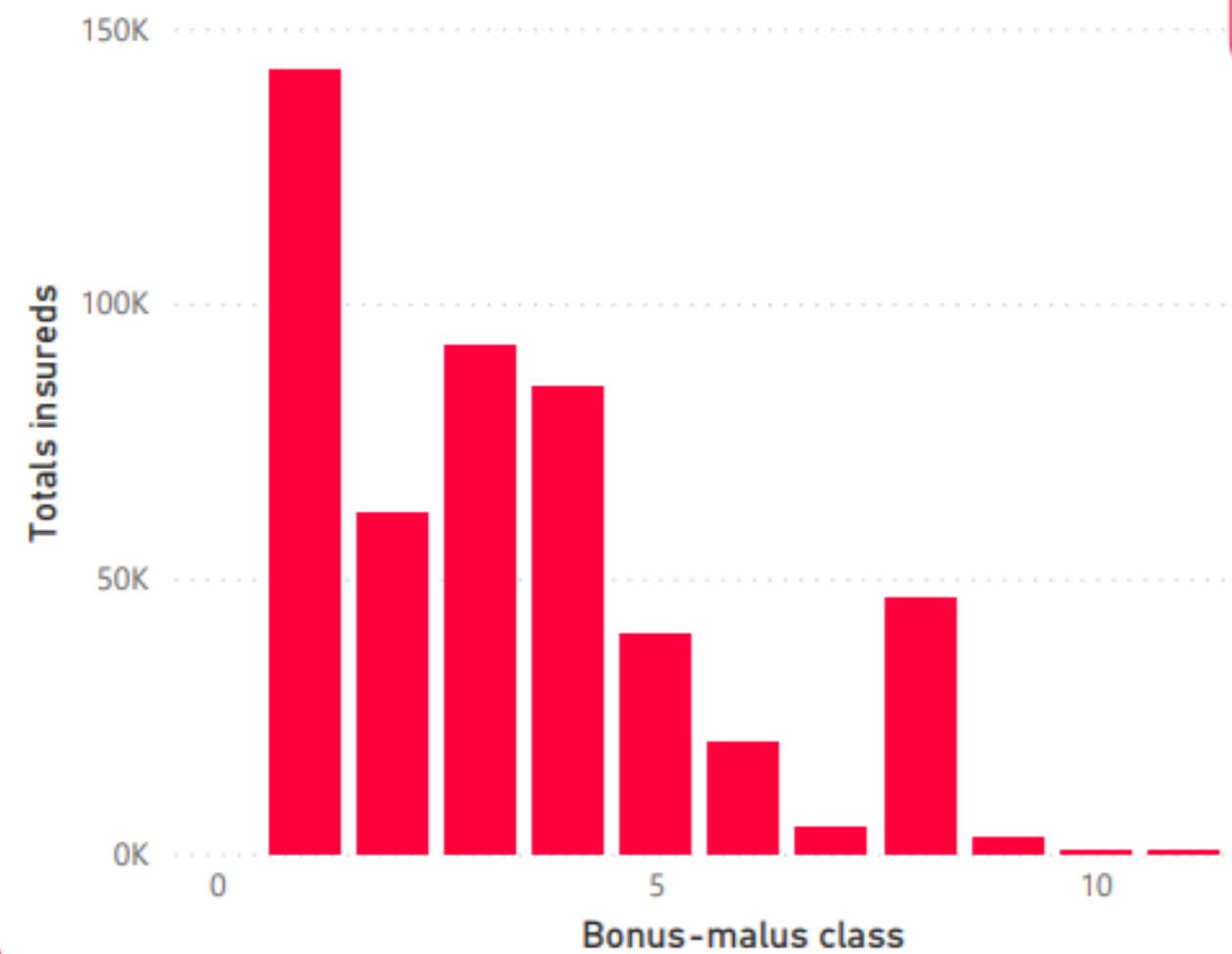
Choose the Bonus-Malus class:

classeBonusMalus

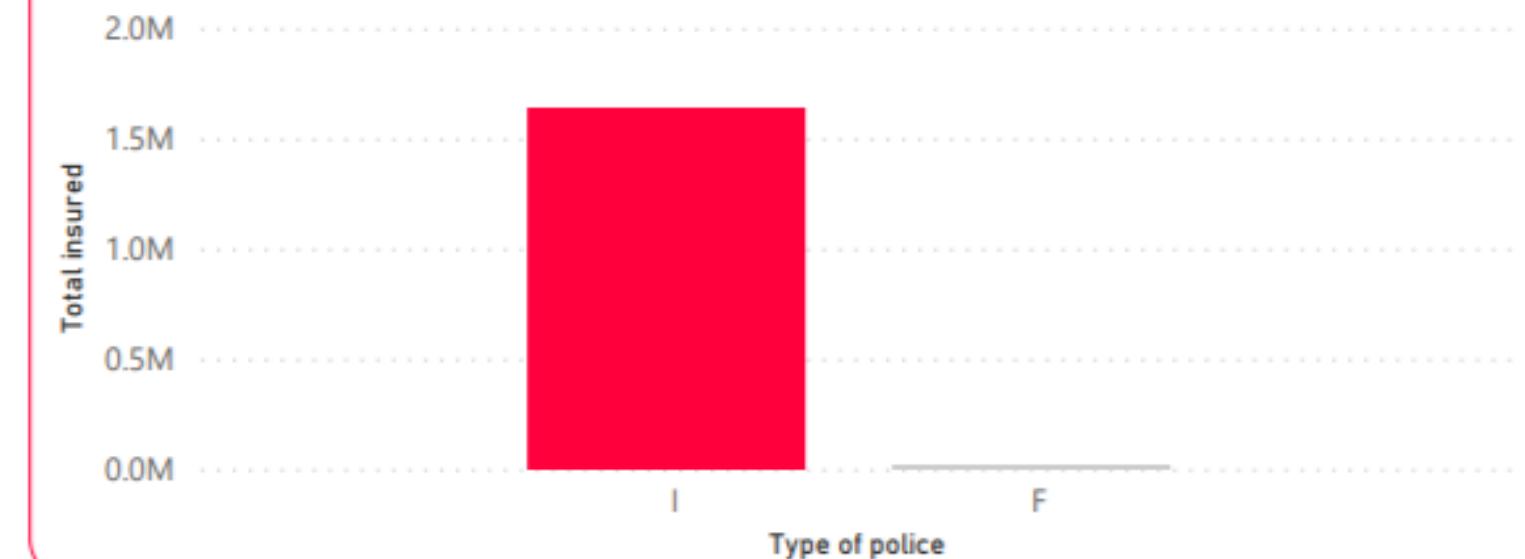
All

497.87K
Totals insurers

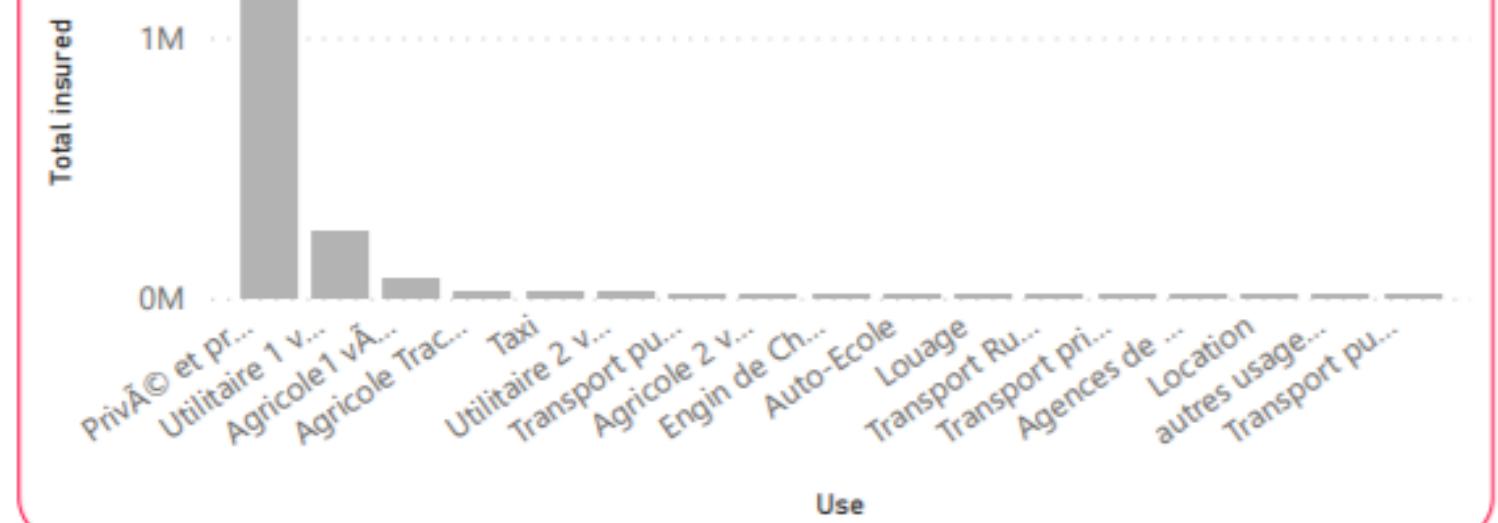
Totals insureds by Bonus-malus class



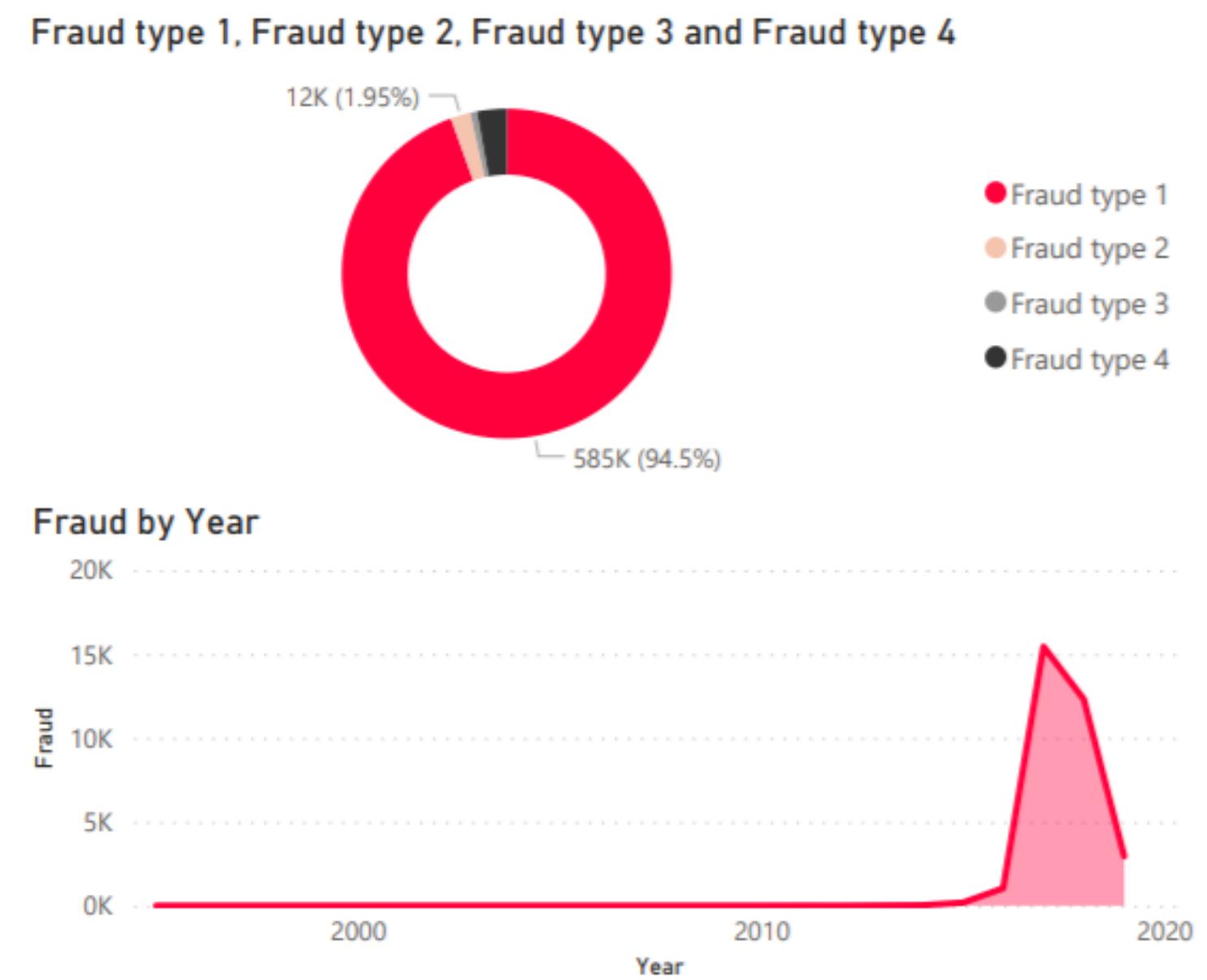
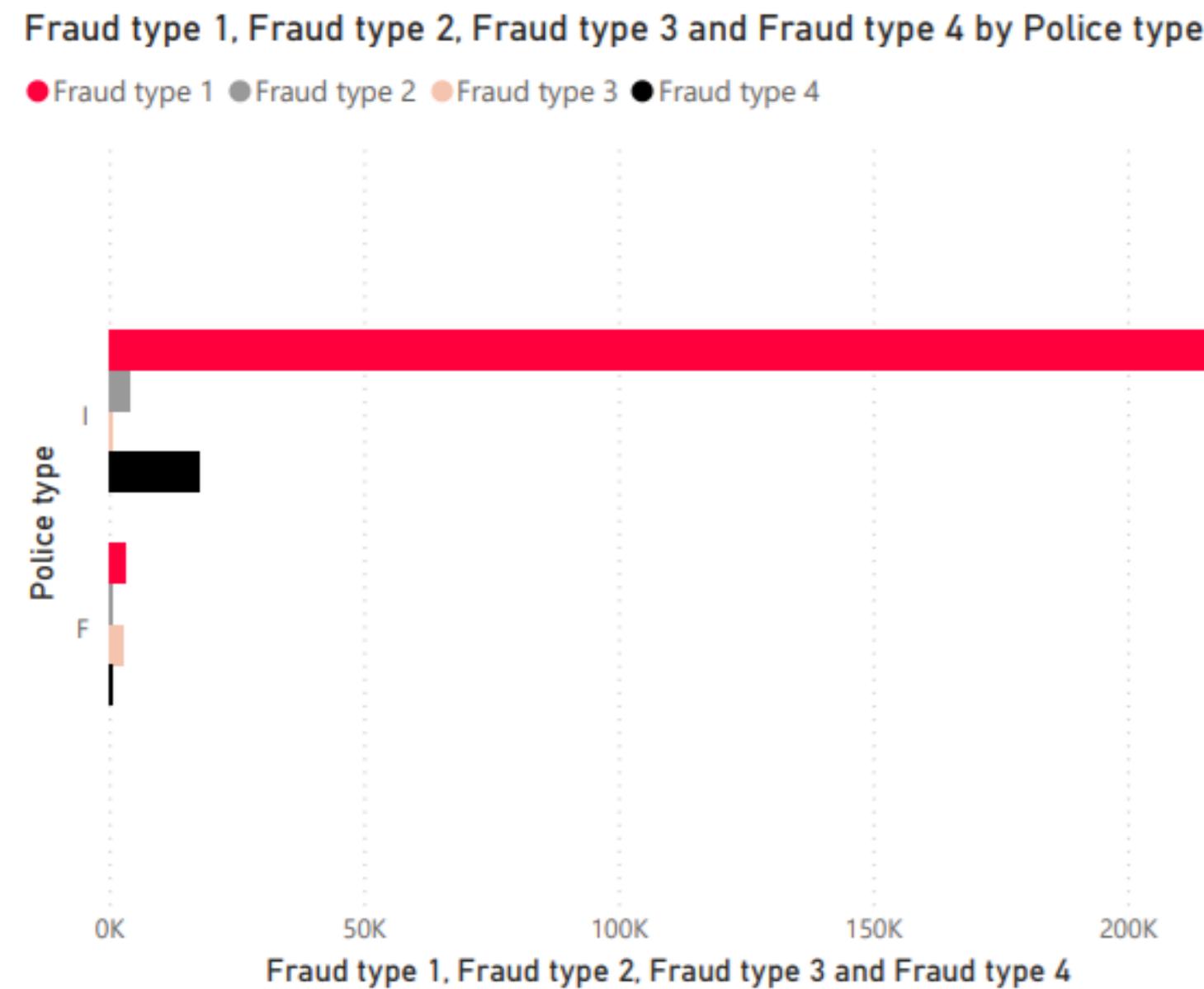
Total insured by Type of police



Total insured by Use



Frauds



Accidents

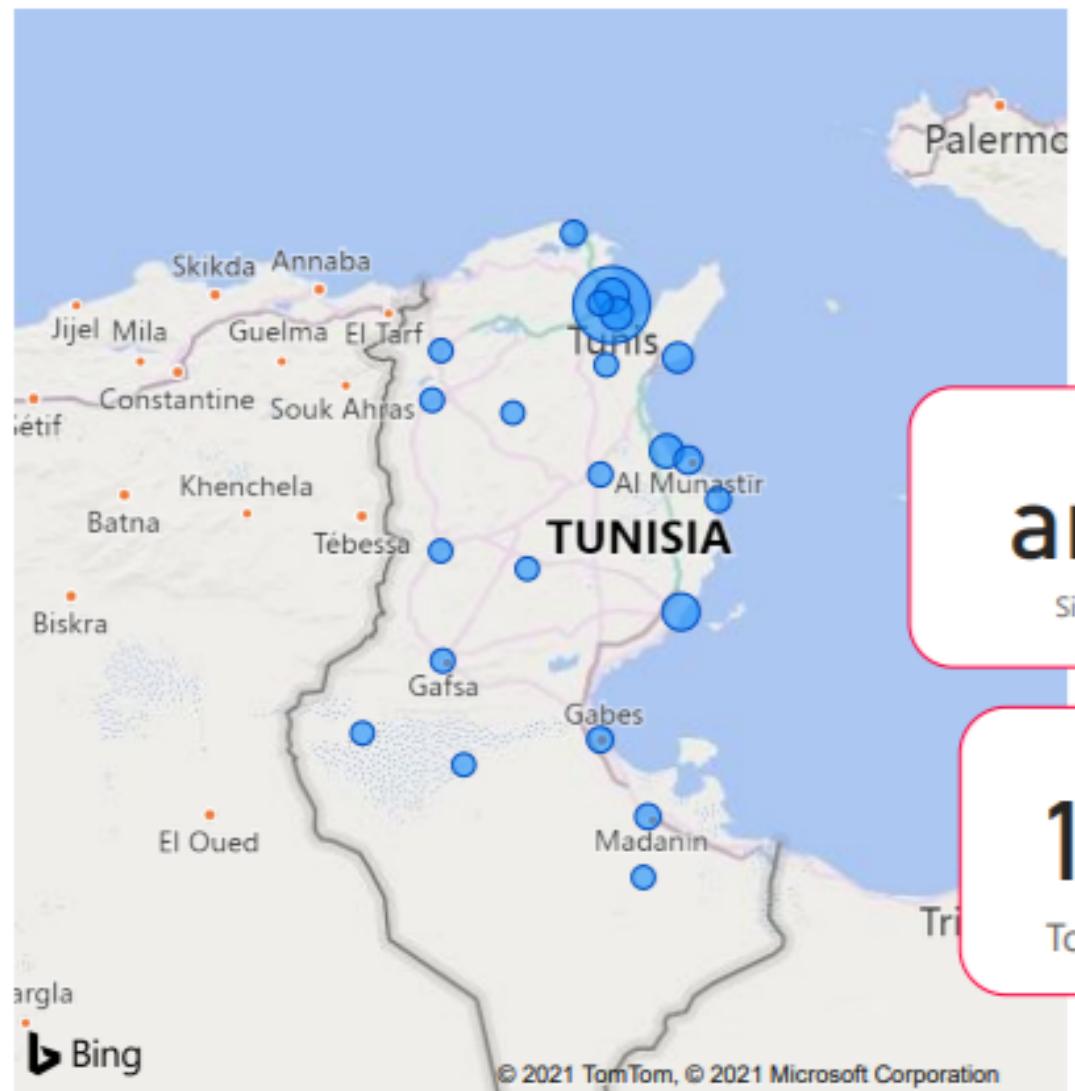
Choose the location:

Sinistre location	Total sinistre
tunis	77553
sfax	20623
ariana	17108
sousse	15174
ben arous	13811
nabeul	10597
monastir	7064
gabes	4308
medenine	3605
bizerte	3404
manouba	2927
mahdia	2560
kairouan	2128
beja	1734
gafsa	1471
kef	1319
zaghouan	1004
Total	190275

Sinistre location

All

Total sinistre by Sinistre location



The weather

10.0

Visibility

10.0

Humidity

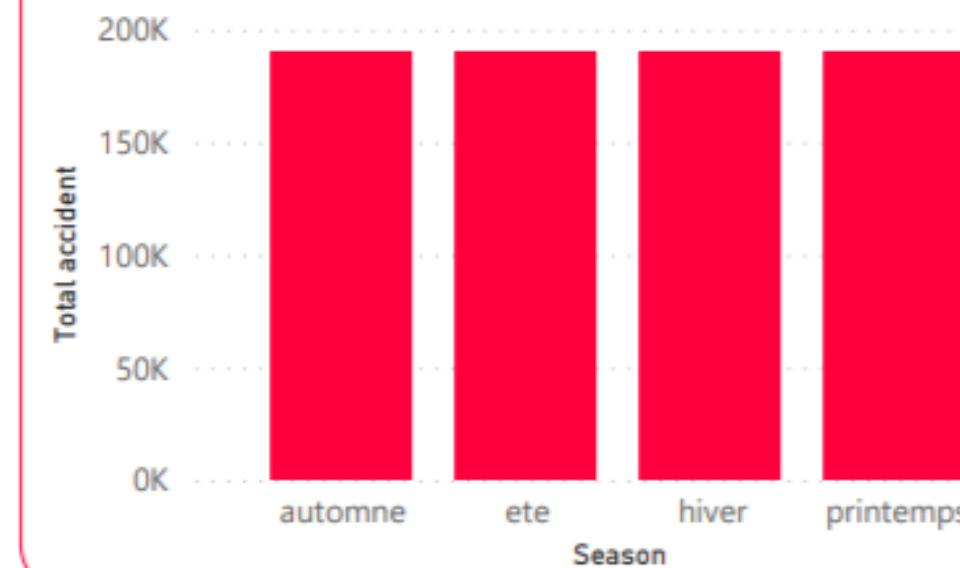
ariana

Sinistre location

190K

Totals sinistre

Total accident by Season





So now let's live this
experience together!



Thank you for your attention

By MNT

#MadeInTunisia