GDPR Data analysis with Power BI

Analyzing GDPR Fines 2018 – 2020 Data with Power BI

Chinh X. Mai

Contents

Case description

GDPR

The **General Data Protection Regulation** (EU) (**GDPR**) is a regulation in EU law on data protection and privacy in the European Union (EU) and the European Economic Area (EEA).

Data

The data includes fines that have been published on the PrivacyAffairs website as of 23.01.2021. It contains the article(s) breached, the type of violation, the fine amount, which country the infraction took place and the authority that handed the fine. Moreover, the organization and source document are also included.

Motivation

Understanding the these fines help prevent future violations and supports the training of the data users to adhere to the GDPR.

Data analytics pipeline

Importing → Cleaning & Transforming → Analyzing & Visualizing → Dashboarding → Communicating
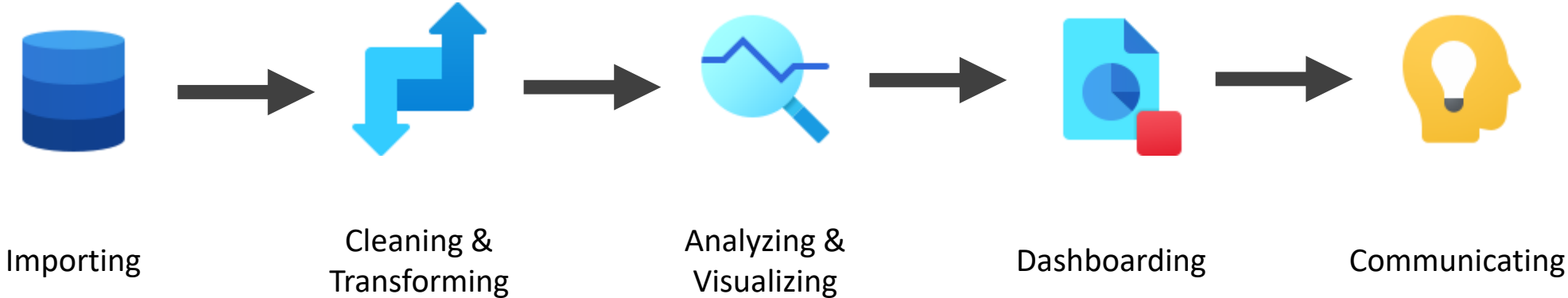
Questions

Examples of question for insight
- Which articles were violated the most?
- What are the most frequent reasons for violations?
- How do fining amounts change over time?
- Which countries have the highest amount of total fine?

Examples of question for design
- Which dimensions should the data be separated?
- Should it be a functional or Explorative dashboard?

Data description

Data source

The data is downloaded from Kaggle, the original page can be accessed by the following link: **GDPR Fines 2018-2020 Updated 23.01.2021**

Data description

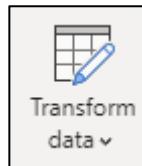The set contains 11 columns and 510 rows, details are given below

id	Index column, *numeric*	**date**	Date when the case took place, *date*
picture	Links to picture of the page, *text*	**org_fined**	Names of the violating organization, *text*
country	Country names, *text*	**articleViolated**	Information of the violated articles, *text*
price	Fine amount, *numeric*	**type**	Reasons of the fine, *text*
authority	Athority issuing the fine, *text*	**source**	Links to the original article, *text*
summary	Summary of the contents , *text*		

Data transformation

Importing → Validating → Cleaning & Transforming → Separting & Connecting

Exploratory Data Analysis
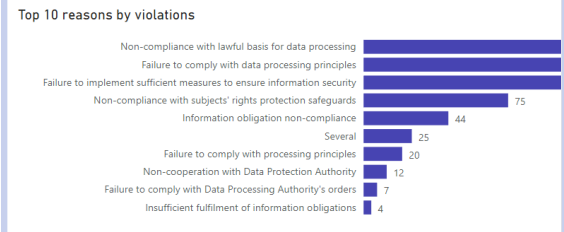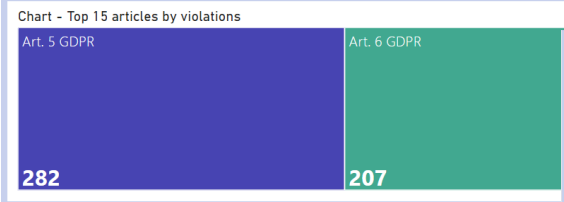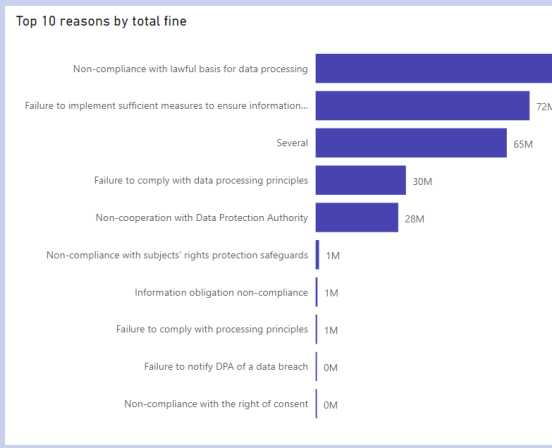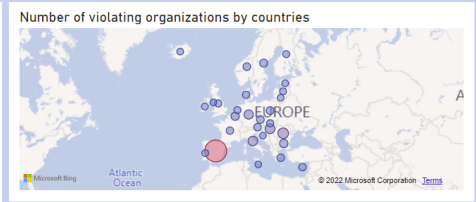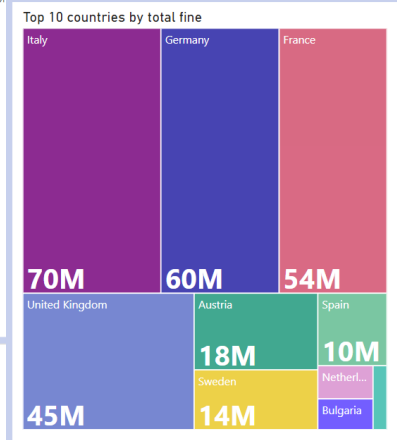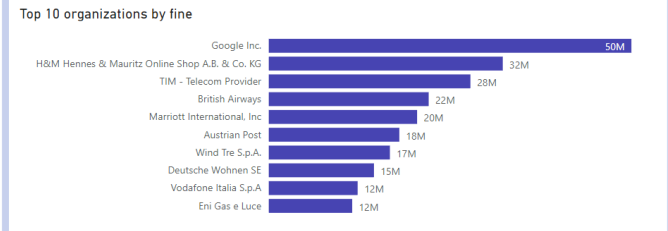
Dashboard design

Factors to consider

- Who is the audience? – **Data analyst team**
- What do they need? – **Identify frequent violations**
- How can a dashboard help? – **Visualizations and interactions**

Dashboard choice

- Should it be a functional or explorative dashboard?
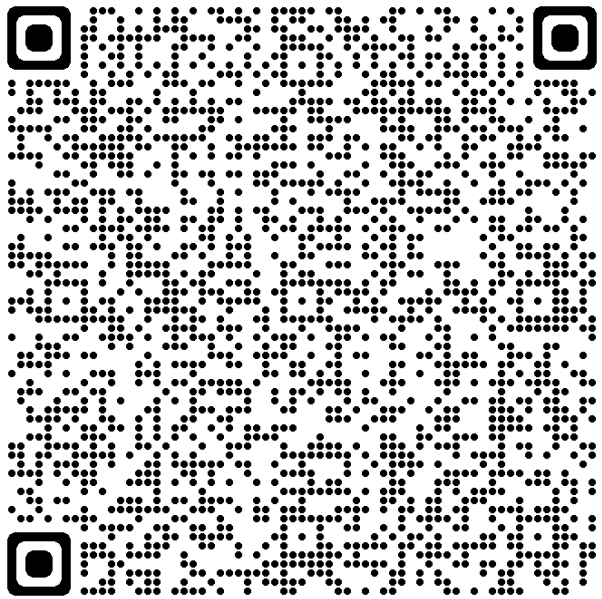 – **Better be a balance between the two**

Final deliverables

Shortened link for web version:
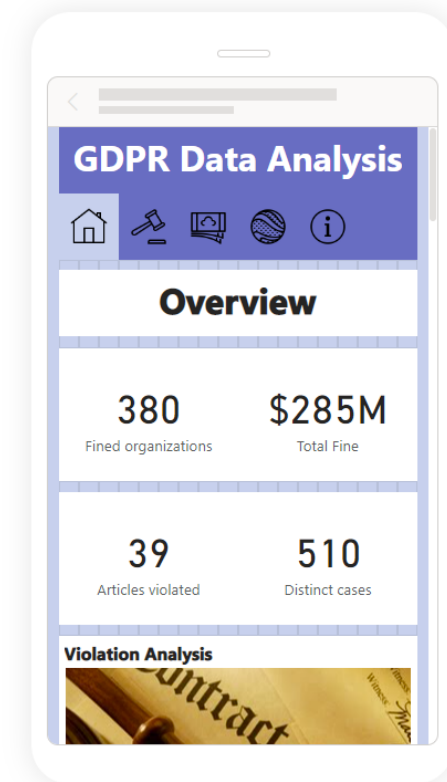**https://bit.ly/3JWgZi9**

Power BI

QR code for mobile dashboard

Final deliverables

Web version

Mobile version

Executive summary

Objectives

- Analyze GDPR data to understand the violating and fining patterns
- Importing, cleaning, transforming, visualizing and carrying out simple data analysis
- Generate a dashboard for exploration using Power BI

Results

- A dashboard in Power BI that is functional on both big screen and mobile
- Facilitate and encourage user engagement