












# GDPR Data analysis with Power BI

Analyzing GDPR Fines 2018 – 2020 Data with Power BI





# Contents

-  Case description
-  Data analytics pipeline
-  Questions
-  Data description
-  Data transformation
-  Exploratory data analysis
-  Dashboard design
-  Final deliverables
-  Executive summary



# 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

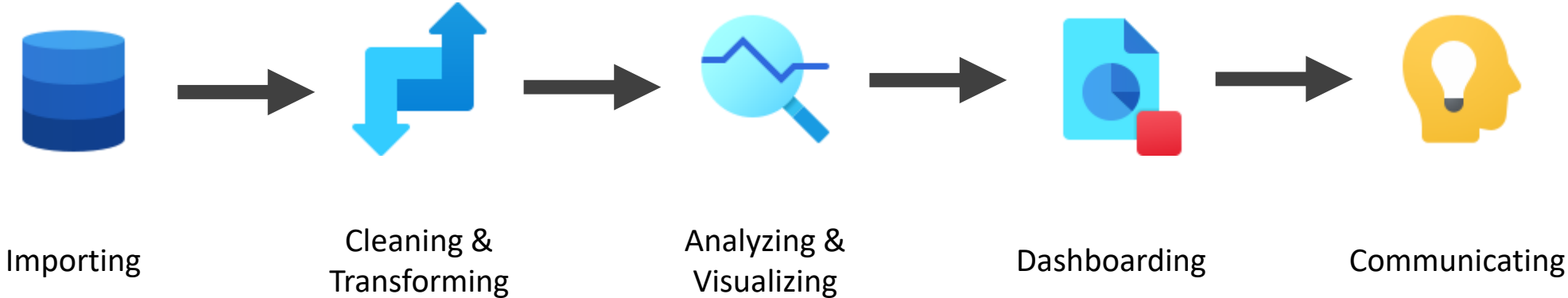
The data includes fines that have been published on the PrivacyAffairs website as of 23.01.2021. It contains the article(s) breached, the type of violation, the fine amount, which country the infraction took place and the authority that handed the fine. Moreover, the organization and source document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

## Motivation

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



# Data analytics pipeline





# 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

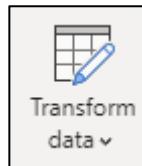
The set contains 11 columns and 510 rows, details are given below

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



# Data transformation

Importing



### Validating

Analysis - Power Query Editor

Transform Add Column View Tools

- Monospaced
- Column distribution
- Show whitespace
- Column profile
- Column quality

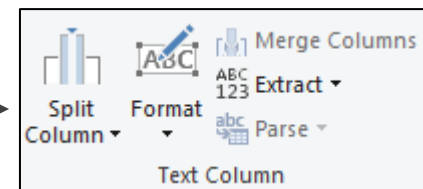
Data Preview

id	articleViolated	type
1	Art. 28 GDPR	Non-compliance with lawful basis for data processing
2	Art. 12 GDPR	Information obligation non-compliance

Column statistics

Count	Value distribution
946	Art. 5 GDPR
0	Art. 6 GDPR
0	Art. 32 GDPR
39	Art. 13 GDPR
15	Art. 15 GDPR
0	Art. 12 GDPR
0	Art. 14 GDPR
0	Art. 33 GDPR
0	Art. 58 GDPR
0	Art. 25 GDPR
0	Art. 21 GDPR

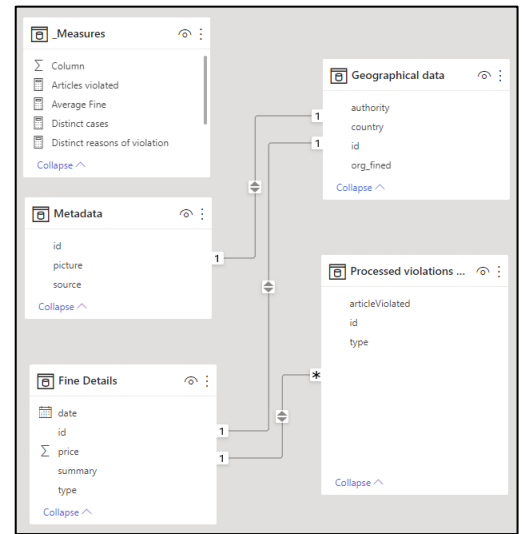
Cleaning & Transforming



- #### APPLIED STEPS
- Source
  - Promoted Headers
  - Changed Type
  - Replaced Value
  - Replaced Value1
  - Replaced Value2
  - Replaced Value3
  - Replaced Value4
  - Replaced Value5
  - Replaced Value6
  - Replaced Value7
  - Replaced Value8
  - Replaced Value9
  - Replaced Value10

Separating & Connecting

- Metadata
- Geographical data
- Fine Details
- Processed violations count

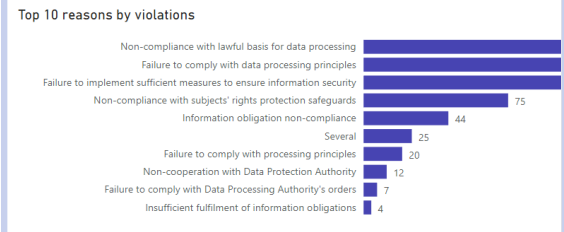
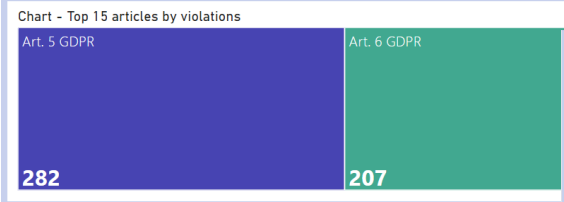




# Exploratory Data Analysis

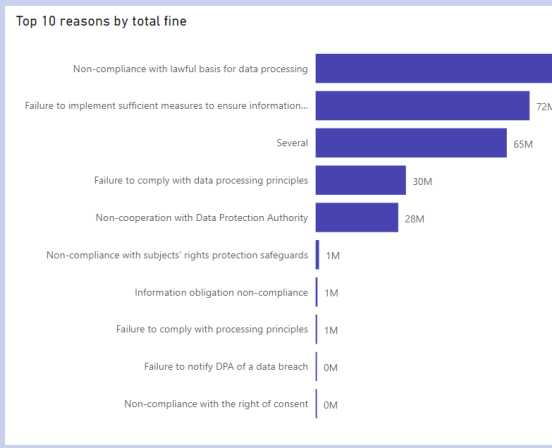
## Violation analysis

26 Distinct reasons of violation    946 Total Violations    39 Articles violated    510 Distinct cases



## Fine analysis

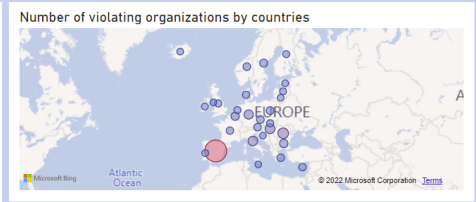
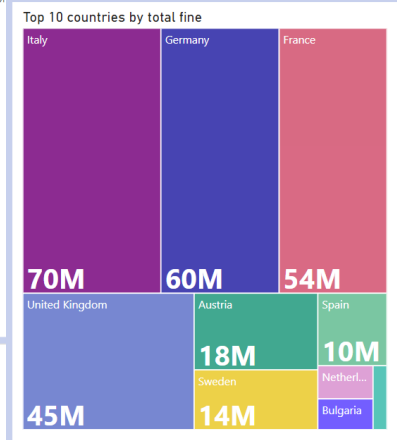
26 Distinct reasons of violation    380 Number of fined organizations    \$559K Average Fine    \$285M Total Fine



## Country analysis

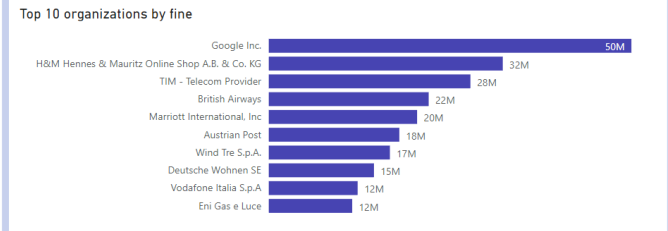
Austrian Data Protection Authority (DSB) Fining authority

Organization name: All  
Country: All



380 Number of organizations

\$285M Total Fine







# 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**

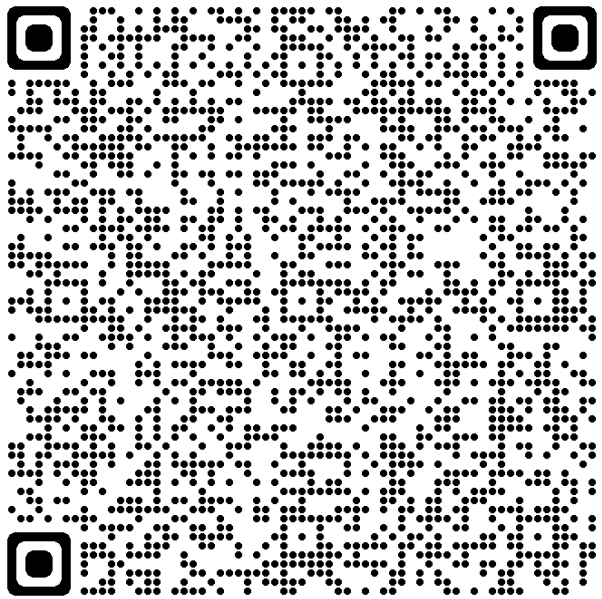
## Dashboard choice

- 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

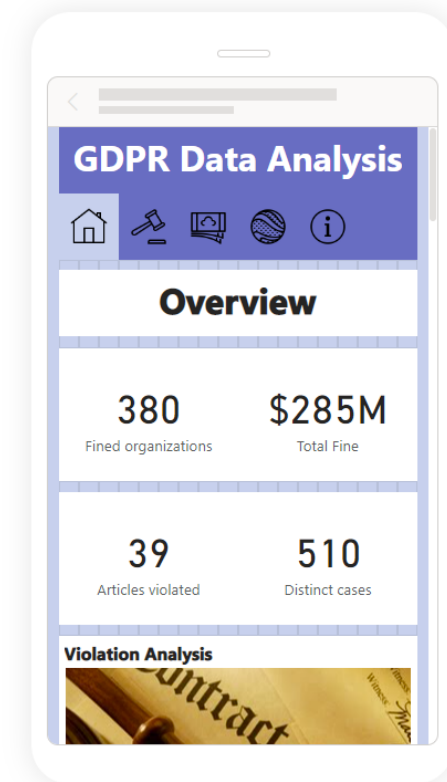
QR code for mobile dashboard



# 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

