Regulations & Laws

# Data Exploration Problem, 250pts

Technical regulations and standards aim to protect human safety, health, the environment, national security, etc. The requirements may concern the features or quality of a product and the procedures for testing, certification, labeling, etc. They relate to issues such as product development, production, packaging, storage, distribution and marketing.

In this challenge, you’ll experiment with the idea that the more often a word or term appears in the body of a technical regulation text, the more semantically significant that the word could be.

As a data scientist, you must analyze the text of the Regulation 2016/679 of the European Parliament (provided as a txt file) and provide keywords which could provide an idea of what is the text about. A summary of the text to enhance the understanding of the regulation topic will get you extra points!

OBJECTIVE:

Analyze the text of the Regulation 2016/679 to provide keywords that could yield to a general understanding of the regulation. For extra points, a summary of the text could be provided to improve its comprehension.

DELIVERABLES:

Presentation (250pts). You will present your findings to a group of judges. They will evaluate your results in terms of how related the key words and the summary are when compared to the text topic.

It will be assessed the originality, technique used to find the words and how a summary could enhance the understanding of the topic.

You will have 5 minutes for your presentation.