Case 1

Consider the following data set: https://github.com/masterfloss/datapolitics. Explore the dataset in the context of a Data Science Project.

Follow a Crisp-DM process that includes the following steps:

- Business understanding
- Data Understanding
- Data Preparation
- Modelling
- Evaluation
- Deployment

Other dimensions may be considered like Organization, Scheduling, Techniques and Risk. How do you expect to consider those dimensions?

To solve this case, answer the following questions.

1. Business understanding. Refer to a context where your data science approach using those data may be helpful.
2. Data Understanding. Create a data dictionary.
3. Data Preparation. What type of data preparation is needed.
4. Modelling. What techniques may be used? What their purposes?
5. Evaluation. Here is analyzed the value for "business".
6. Deployment. What forms of deployment do you suggest?
7. How to organize, schedule, manage risk and techniques selection.