What should I expect the data format to be?

Each sample in the dataset has the following information:

* Interactions between doctor and patient
* Identifier for doctor
* Identifier for patient (optional)

Columns:

* Doctor\_name
* Doctor\_address
* Patient\_name (optional)
* Problem\_type (eg, Fever, stomach related, etc)
* Interactions / Reviews (text)
* Rating