Earlier models have only a description.

To ask :

1. Activities?

Trained on training data

Tasks:

1. Preprocess data:(Take help from previous code for description)
   1. Cleaning of data scratch. (handling missing values, coveting category data to numeric data).
   2. Preprocessing of text data description.
      1. Remove outlier(description with one letter eg. “desc”)
      2. Remove location from the description
         1. First, create a list of locations present in the data
         2. Then go through each data and remove the location
         3. Remove words like calfresh or any other word which doesn't make sense.
         4. Remove stopWords(The, a then, he, she)(inBuild function or check for earlier code)
   3. Preprocess columns other than description(category col to numeric).
   4. Add Co variant(co-relation) plot and frequency plots
2. Develop a way to aggregate the activities to classify the partnership( not the activities themselves)
3. Utilize partnership\_data as training data

Topic Modelling :

Clustering :

We need to add more data cols to the model.

We need to check how we give model text data plus other columns.

Notes: Done by Previous Members :

Classification of High And Low-level Meetings:

High-Level Meeting: Has more meetings, better work mentioned in the description

Low-Level Meeting: Has fewer meetings, less work mentioned in the description