Original location of the dataset (ideally a URL to the original data): http://visualizing.org/datasets/new-cases-cancer-us License information (e.g. whether this dataset can be shared publicly): This is a public dataset. (source:CDC/NCHS) Type of data in the dataset (e.g. numerical, categorical, unstructured, and so on): numerical and categorical, structured. Why you chose this dataset (3 to 5 sentences): This is a dataset about how new cases of cancer are distributed over cancer type, ethnicity and time. I am interested what kind of insights can be drawn from it. I also think it is a good dataset to visualize since it is a time series data with numerical and categorical features. Link to the dataset in your own repository: https://github.com/stuncyilmaz/stuncyilmaz.github.io/tree/master/data_visualizations/project