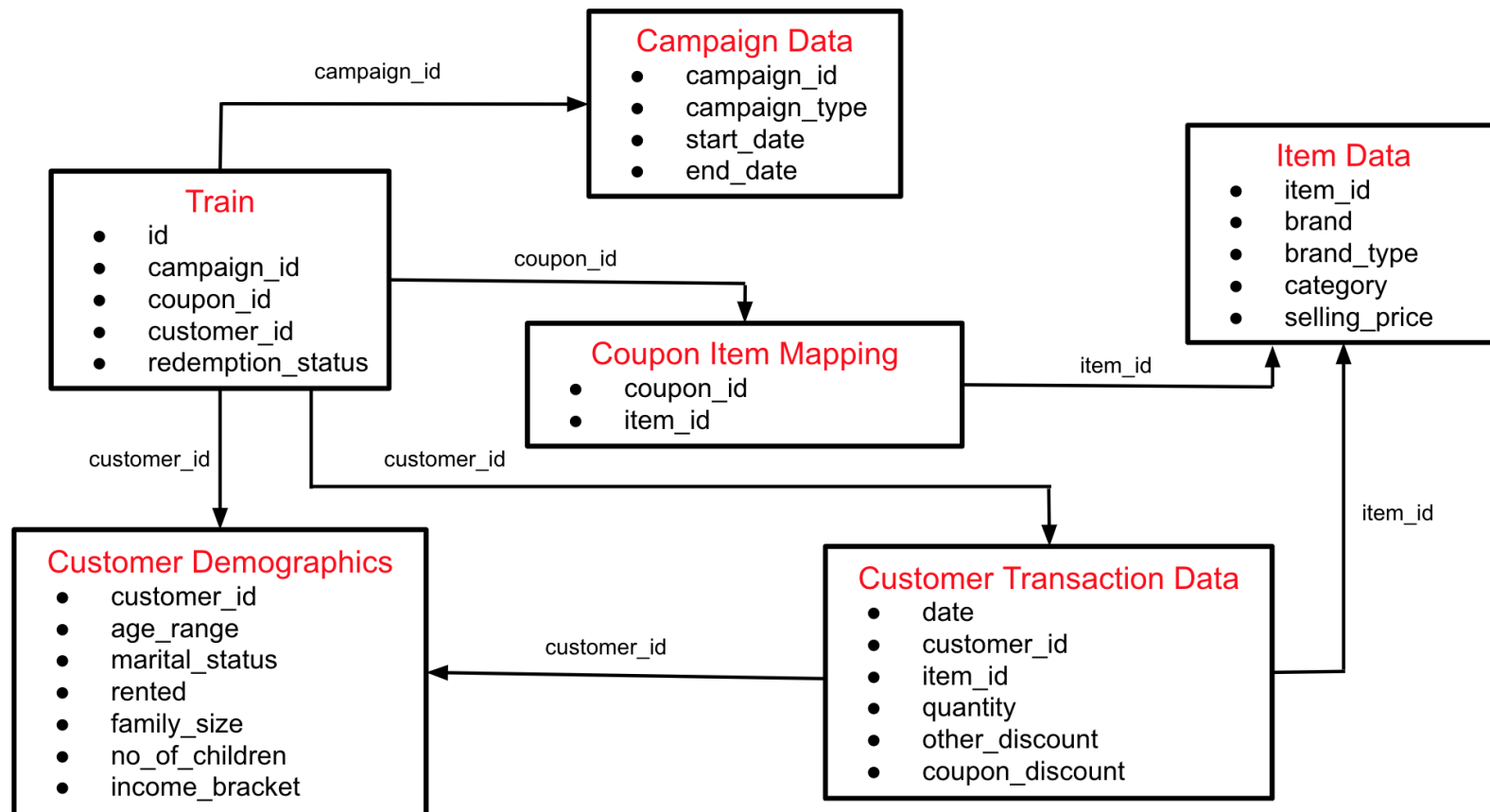


Data Analysis and Summarization using Pandas

You are given the following data as 6 separate csv files:



Perform the following:

1. Initial data analysis for all datasets individually.
2. A report containing a summary of your understanding of the data, important fields, plan for next steps, etc.
3. Data cleaning and transformations.
4. Exploratory data analysis - individual and as a whole.
5. Merge all datasets into one meaningful and summarized dataset.
6. A report explaining your thought process for the chosen method(s) for merging all datasets and the approach(es) taken to summarize everything into one dataset, also potential uses of your summarized dataset.

Note:

- a) Try to do everything using Pandas only.
- b) Other libraries should not be used unless there is a strong justification to do so.
- c) All your code, code comments, output screenshots, documentation, reports and diagrams should be in a single ipynb file uploaded to your Github repo and the link to this file in Github should be submitted as your project in Google Classroom.