T5 - Data analysis techniques and methodologies

Merging on Dataframes Columns

We can merge Dataframes N:1 and N:N

pandas.merge(<Dataframe_1>, ..., <Dataframe_n>) -> Looks for strictly coincidences index and labels

pandas.merge(<Dataframes>, on= <ColumnLabels>) -> Label exact coincidence values, not indexes.

pandas.merge(<dfs>, on = <CoLabels>, how= {'inner', 'right', 'left', 'outer' }) -> Order to merge:

inner: for labels, after indexes df1, …, after indexes dfn. Default value for merge

right: for Colabels, after labels dfn, …, after labels df1. NaN not permitted on right dfs labels

left: for CoLabels, after labels df1, …, after labels dfn. NaN not permitted on left dfs labels

outer: same as inned, but permits NaN for any non combination.

pandas.merge(…., suffixes=<suffix list for labels not in CoLabels>)

Merging on DataFrames Indexes

Merge index to index -> **left_index = True, right_index = True**

Merge label with index -> left_on = < list of labels>, right_index = True

|-> right_on=< list of labels>, left_index = True

Joining Dataframes with same indexes.