- 第8章、连接、联合、重塑
  - 8.1 分层索引
    - 元组作索引
    - 元组作字典键
    - 二维数组作索引
    - pd.MultiIndex.from_arrays/from_tuples/from_product([], [])
    - 取值
      - s
        - 外层
          - s.loc[['外层索引列表']]
          - s.loc[索引切片]
        - 内层
          - s.loc[:, ['内层索引列表']]
        - 取值
          - s.loc['外层', '内层']
      - df
        - 外层
          - df.loc[:, ['外层列列表']]
          - df.loc[['外层行列表'], :]
        - 内层
          - df.loc[:, pd.IndexSlice[:, ['内层列列表']]]
          - df.loc[pd.IndexSlice[:, ['内层行列表']], :]
        - 取值
          - df.loc[:, ('外层列', '内层列')]
          - df.loc[('外层行', '内层行'), :]
          - df.loc[('外层行', '内层行'), ('外层列', '内层列')]
    - 删除层级索引
      - df.index.droplevel(level=)
      - df.columns.droplevel(level=)
    - 多层索引Series转DataFrame
      - s.unstack()
      - 逆向 df.stack()
    - 8.1.1 交换层级顺序、层级排序
      - 交换层级顺序 df.swaplevel(i=-2, j=-1, axis=0)
      - 按照某一层级进行排序 df.sort_index(axis=0, level=)
    - 8.1.2 层级聚合 聚合函数中使用level参数 df.sum(level=列names, axis=1)
    - 8.1.3 使用列构造层级索引
      - df.set_index([列列表])
      - df.reset_index()
  - 8.2 合并数据集
    - 8.2.1 数据库风格 pd.merge(left, right, how='inner', on, left_on, right_on, left_index, right_index, sort=True, suffixes=('_x', '_y'), indicator=False指明数据的来源有left_only、right_only、both)
    - 8.2.2 根据索引合并 df1.join([df2, df3, df4...], on=df1的列作为索引, how='left')
    - 8.2.3 沿轴连接
      - np.concatenate(arr序列, axis=0)
      - pd.concat([objs序列], axis=0, join='outer' / 'inner', ignore_index=False, keys=[序列]沿着axis增加最外层索引变成多层索引, levels=[序列]构建多层索引, names=[列表]多层索引命名)
    - 8.2.4 利用另一个对象补充缺失值
      - np.where(pd.isnull(a), b, a)
      - df1.combine_first(df2) 使用df2的数据填充df1的NaN值
  - 8.3 重塑、透视
    - 8.3.1 重塑 DataFrame转Series df.stack(level=-1)
    - 8.3.2 透视
      - pd.pivot_table(data=df, values=聚合的列, index=列的值变成索引, columns=列的值变成列, aggfunc='mean', fill_value=None)
      - df.pivot(index=列的值变成索引, columns=列的值变成列, values=列的值用来填充结果) 等价于
        - df.set_index([pivot中的参数index, columns]).unstack(pivot中的参数columns)
        - df.pivot('date', 'item') == df.set_index(['date', 'item'].unstack('item'))
    - 8.3.3 透视逆操作 pd.melt(frame, id_vars=列的值作为分组依据, value_vars=结果中variable列的值，如果不指定，值为df去除id_vars的所有列名, var_name='variable', value_name='value') pd.melt(df, ['key']).pivot('key', 'variable', 'value').reset_index()