It's a dataset of 2 year EURUSD basis and its possible contributors.

1. Columns in the sheet (target variables and features)
   - The **target variable** is the 'd(basis)' column which indicates the changes of basis.
   - The columns behind the 'd(basis)' column are the features.
   - **'d(basis)-1'** denotes change of basis during last period.
   - **'$index'** denotes Bloomberg US Dollar index which shows the strength of US Dollar.
   - **'VIX'** denotes CME global volatility index.
   - **'TED'** denotes ted spread (3m US libor-3m treasury yield).
   - **'Euribor-GGB'** denotes 3m Euribor-3m Generic Germany bond yield
   - **'LIBOROIS'** denotes overnight libor/ois spread.
   - **'EuriborEONIA'** denotes 3m Euribor-3m EONIA
   - **'revyankee'** denotes American corporates' issuance of Euro bonds.
   - **'solvency risk'** denotes BAML GFSI Solvency Component.
   - Since this project aims to predicting changes of basis, if in a sample, 'date' is t, 'd(basis)' is the change of basis during period t, other features are values during period t-1.

2. Time span and data frequency
   - Data's time span is from **2005/4/26 to 2017/4/10**.
   - In the file, there are two sheets—'weekly data' and 'daily data'. In 'weekly data' and 'weekly data2', there are 622 samples totally. In 'daily data', there are 2958 samples totally.
   - Compared to 'weekly data', 'weekly data2' has an extra 'd(basis)-1' column.
   - There isn't 'revyankee' column in 'daily data' due to lack of statistics.

3. Data processing
   In daily data, different kinds of statistics have slightly different dates, and only the data share the same dates are kept.