

Assumption:

- QuantDataSet and TextDataSet are for classification
- Data is loaded from a file path
- TimeSeriesDataSet here is supposed to track real-time value with a time step

ABC_DataSet

- location : string
- filename : string
- description : string
- size : float

+ABC_DataSet(location:string, filename:string, description:string, size:float):
+insertData():void
+deleteData():void
+printData():void
+getLocation():string
+getFilename():string
+getDescription():string
+getSize():float
+getColumn_name():string
+getRow_name():string

Explanation:

Design Main ABC_DataSet class for general data operation purposes and we extended it with various subclasses (Quant, TimeSeries, Text) for specific data characteristics.

- **Quant**: Quantitative data processing and data plotting.
- **TimeSeries**: Retrieve attributes from Time Series data and data plotting.
- **Text**: Text data clustering and advanced text data analysis.

extends

extends

extends

QuantDataSet

- super(ABC_DataSet)
- label: string
- quantType:string
- frequency:int

+getStatistics():float
+removeOutliers():void
+getAggregation():float
+getBarplot():void

TimeSeriesDataSet

- super(ABC_DataSet)
- time: timestamp
- step: float
- coordinates:tuple

+getTime_range(): void
+getTimestampData():void
+getHistogram():void
+getStatistics():float

TextDataSet

- super(ABC_DataSet)
- corpus:int
- language:string
- label:string

+letter_standardization():void
+textClustering():void
+textCategorization():void
+wordPartOfSpeech():void